Structure and Evolution of the Actin Gene Family in *Arabidopsis thaliana*

John M. McDowell,* Shurong Huang,† Elizabeth C. McKinney,‡ Yong-Qiang An‡ and Richard B. Meagher‡

**Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599, †Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720 and ‡Department of Genetics, University of Georgia, Athens, Georgia 30602*

Manuscript received August 10, 1995
Accepted for publication October 27, 1995

ABSTRACT

Higher plants contain families of actin-encoding genes that are divergent and differentially expressed. Progress in understanding the functions and evolution of plant actins has been hindered by the large size of the actin gene families. In this study, we characterized the structure and evolution of the actin gene family in *Arabidopsis thaliana*. DNA blot analyses with gene-specific probes suggested that all 10 of the Arabidopsis actin gene family members have been isolated and established that Arabidopsis has a much simpler actin gene family than other plants that have been examined. Phylogenetic analyses suggested that the Arabidopsis gene family contains at least two ancient classes of genes that diverged early in land plant evolution and may have separated vegetative from reproductive actins. Subsequent divergence produced a total of six distinct subclasses of actin, and five showed a distinct pattern of tissue specific expression. The concordance of expression patterns with the phylogenetic structure is discussed. These subclasses appear to be evolving independently, as no evidence of gene conversion was found. The Arabidopsis actin proteins have an unusually large number of nonconservative amino acid substitutions, which mapped to the surface of the actin molecule, and should effect protein-protein interactions.

ACTIN is a ubiquitous component of the plant cytoskeleton and participates in a number of important subcellular processes (reviewed in STAIGER and SCHLIWA 1987; LLOYD 1988, 1991; STAIGER and LLOYD 1991; MEAGHER and WILLIAMSON 1994). The microfilament cytoskeleton is involved in cell division plane localization (ELEFTHERIOU and PALEVITZ 1992; LLOYD and TRAAS 1988; MINEYUKI and PALEVITZ 1990), cell elongation, and cell shape determination (KOBAYASHI *et al.* 1987; SEAGULL 1990; WERNICKE and JUNG 1991). These functions are of central importance to plant development, because the constraint of rigid cell walls dictates that morphogenesis in plants occurs largely through asymmetric cell division and expansion. In tip-growing pollen tubes, membrane vesicles and generative cells move by actin-based motility (reviewed in MASCARENHAS 1993; RUSSELL 1993). Organellar movement (*e.g.*, light-responsive chloroplast orientation) and cytoplasmic streaming are thought to be driven by an acto-myosin motor (WILLIAMSON 1993). Colocalization of polyribosomes with microfilaments has also been reported (ITO *et al.* 1994), indicating that the actin cytoskeleton influences polysome intracellular distribution.

The diversity of these functional roles is paralleled by the diversity within plant actin gene families. Although plant actins are very similar to other eukaryotic actins (>83–88% amino acid identity with most animal and fungal actins), they are encoded by gene families that are much more diverse than those in other eukaryotes. Most metazoans contain fewer than 10 functional actin genes, and many fungi and protists contain only one or a few actin genes. Petunia, however, contains >100 actin sequences in its genome (BAIRD and MEAGHER 1987), and other plant species such as soybean, tobacco, potato, rice, and lodgepole pine also appear to have dozens of actin genes (MEAGHER 1991; MEAGHER and WILLIAMSON 1994; THANGAVELU *et al.* 1993). Plant actin genes also exhibit more intraspecific divergence (6–10% nonsynonymous nucleotide substitution within the soybean and rice actin gene families) than do the 500 million-year-old animal muscle and cytoplasmic actin gene subclasses (3–7% nonsynonymous substitution) (HIGHTOWER and MEAGHER 1986). An obvious, and largely unanswered question, is why do plants contain such complex actin gene families?

One possible answer is that plants employ specialized actin genes for the diverse functions described above. Gene duplication, followed by divergence in regulatory regions and/or protein coding sequence, has long been recognized as a potent source of genes with novel functional capabilities (HALDANE 1932; OHNO 1970). For example, warm-blooded vertebrates contain six actin isoforms that have distinct differences in their sequences and patterns of expression (MIWA *et al.* 1991). The genes encoding these isoforms arose from a series of ancient gene duplications that occurred 200–500 mya. The unique expression patterns and primary sequences of these six isoforms have been conserved for >250 million years, and recent experimental evidence

Corresponding author: Richard B. Meagher, Department of Genetics, University of Georgia, Athens, GA 30602.
E-mail: meagher@bscr.uga.edu

(reviewed in HERMAN 1993) suggests that these genes have specialized functions that are required for proper development of their respective cell types. Essentially nothing is known about the roles of different plant actin genes, but the observed diversity and divergence within plant actin gene families suggests that they may have also evolved distinct functional capabilities.

Tissue-specific and developmental expression of actin genes has been observed in several plant species (reviewed in MEAGHER 1991; MEAGHER and WILLIAMSON 1994). In soybean, mRNA levels from the κ, λ, and μ subclasses vary by 100-fold (HIGHTOWER and MEAGHER 1985). The κ actins are expressed in all root cell types except the cap, while the λ actin subclass is expressed only in root protoderm and root cap (MCLEAN *et al.* 1990). In tobacco, the TAc9 gene is expressed in root, leaf, pollen, and stigma. The TAc25 gene, however, is expresssed exclusively in mature pollen (THANGAVELU *et al.* 1993; D. A. BELOSTOTSKY and R. B. MEAGHER, unpublished data). In rice, the RAc2 and RAc3 message levels drop fivefold by 13 days after germination, while the RAc1 and RAc7 message levels remain relatively constant (MCELROY *et al.* 1990).

Preliminary studies of plant actin evolution have suggested the existence of plant actin gene subclasses that are at least as ancient as those in animals. It has been estimated (HIGHTOWER and MEAGHER 1986) that the κ, λ, and μ soybean actin subclasses diverged during the emergence of vascular plants, ~300–500 mya. Based on this observation, and the data discussed above, HIGHTOWER and MEAGHER hypothesized that plant actin gene families contain ancient subclasses that have acquired specialized functions and/or patterns of expression. This hypothesis, when considered in terms of the importance of the actin cytoskeleton in plant growth and development, suggests a potentially significant link between actin gene evolution and macroevolution of plant structures, tissues, and organs (MEAGHER 1994). Unfortunately, this hypothesis has been difficult to address because of the large size of most plant actin gene families. A comprehensive study of the structure, function, and evolution of an actin gene family in even one plant species is currently lacking.

We are using *Arabidopsis thaliana* as a model to address the significance of actin gene multiplicity in plants. In this paper, we present the results of a study in which we thoroughly characterized the Arabidopsis actin gene family, examined structural features of the encoded proteins, and determined the evolutionary relationships of these genes to each other and to actin genes in other eukaryotes. This study, along with recent companion studies of actin gene expression, which are discussed herein, suggest that Arabidopsis contains two ancient actin classes that are partially specialized for reproductive and vegetative function and regulation. These classes diverged further into six subclasses that have evolved significant functional and regulatory differences.

MATERIALS AND METHODS

Overview of strategy for gene isolation: We screened about five genomic equivalents (25,000 clones) from two different Arabidopsis genomic libraries and over 2×10^6 clones from three cDNA libraries with multiple actin probes under low stringency. We then amplified a fragment of coding sequence from each putative actin clone with degenerate actin PCR primers, which annealed to highly divergent actin genes. We obtained ~200 nucleotides of diagnostic sequence from each PCR product and determined whether the clone was unique or had already been isolated. Clones with unique actin sequences were purified to homogeneity. The screening continued until every gene had been isolated four times. Additional details of the cloning of *ACT 1* and *3* (AN *et al.* 1996), *ACT2* and *8* (AN *et al.* 1996), *ACT4* and *12* (HUANG *et al.* 1996), *ACT7* (J. MCDOWELL, Y.-Q. AN, E. C. MCKINNEY, S. HUANG and R. B. MEAGHER, unpublished data), and *ACT11* (S. HUANG, Y.-Q AN, J. MCDOWELL, E. C. MCKINNEY and R. B. MEAGHER, unpublished results) are discussed in forthcoming manuscripts.

Library screening: The following *A. thaliana* cv. Columbia libraries were screened as described (NAGAO *et al.* 1981): genomic DNA partially digested with Sau3A in λfix (kindly provided by BRIAN HAUGE, Massachusetts General Hospital); randomly sheared genomic DNA in λGEM11 (kindly provided by JOHN MULLIGAN, Stanford University); leaf cDNA in pcDNAII (MEAGER lab); leaf cDNA library in λZap (kindly provided by PEGGY HATFIELD, University of Wisconsin); root cDNA in λgt11 (kindly provided by BRIAN HAUGE). To avoid biasing the hybridization conditions toward a particular subset of actin sequences, we screened duplicate filter imprints of the plaques or colonies on each plate with heterologous hybridization probes from Dictyostelium and Soybean (SAc3) (HIGHTOWER and MEAGHER 1985) or Petunia (PAc1) (MCLEAN *et al.* 1988). Hybridization and wash conditions were as described in BAIRD and MEAGHER (1987), except that the hybridizations were at 52° and the washes at 56°. Clones that hybridized to both actin probes in the primary screening were selected for further analysis with PCR sequencing.

Primers for PCR screening of putative actin clones: To design PCR primers that would anneal to evolutionarily diverse plant actin sequences, an alignment of actin amino acid sequences was used to identify regions that were completely conserved among the known plant, animal, fungal, and protist actins. Based on this information, two nested sets of degenerate oligonucleotides were designed (Figure 1).

PlAc12S: TGYGAYAAYGGNACNGGNATGG
PlAc46S-20: ATGGTNGGNATGGGNCARAA
PlAc245N-20: GTDATNACYTGNCCRTCNGG
PlAc284N: ATRTCNACRTCRCAYTTCATDAT

The above sequences are listed 5′ to 3′. The number refers to the N-terminal codon in each sequence. "S" and "N" designate that the sequence of the oligonucleotide is identical to the sense and nonsense strand of the gene, respectively. The 46S-20 primer was chemically phosphorylated during synthesis for use in the λ exonuclease digestion described below. Oligonucleotides were synthesized at the Molecular Genetics Instrumentation Facility, University of Georgia.

Polymerase chain reactions: The above primers were used to amplify a portion of the actin gene from each clone. Because most of the amplifications were done from unpurified plug stocks containing hundreds of non-actin clones in addition to the actin-containing clones, we used "nested" PCR to increase the specificity and yield of the amplification. For genomic clones, the 12S and 284N primers were utilized in the first amplification. The resulting PCR product was precipitated with 0.6 volumes isopropanol, 0.3 M NaOAc, pH 5.2, to

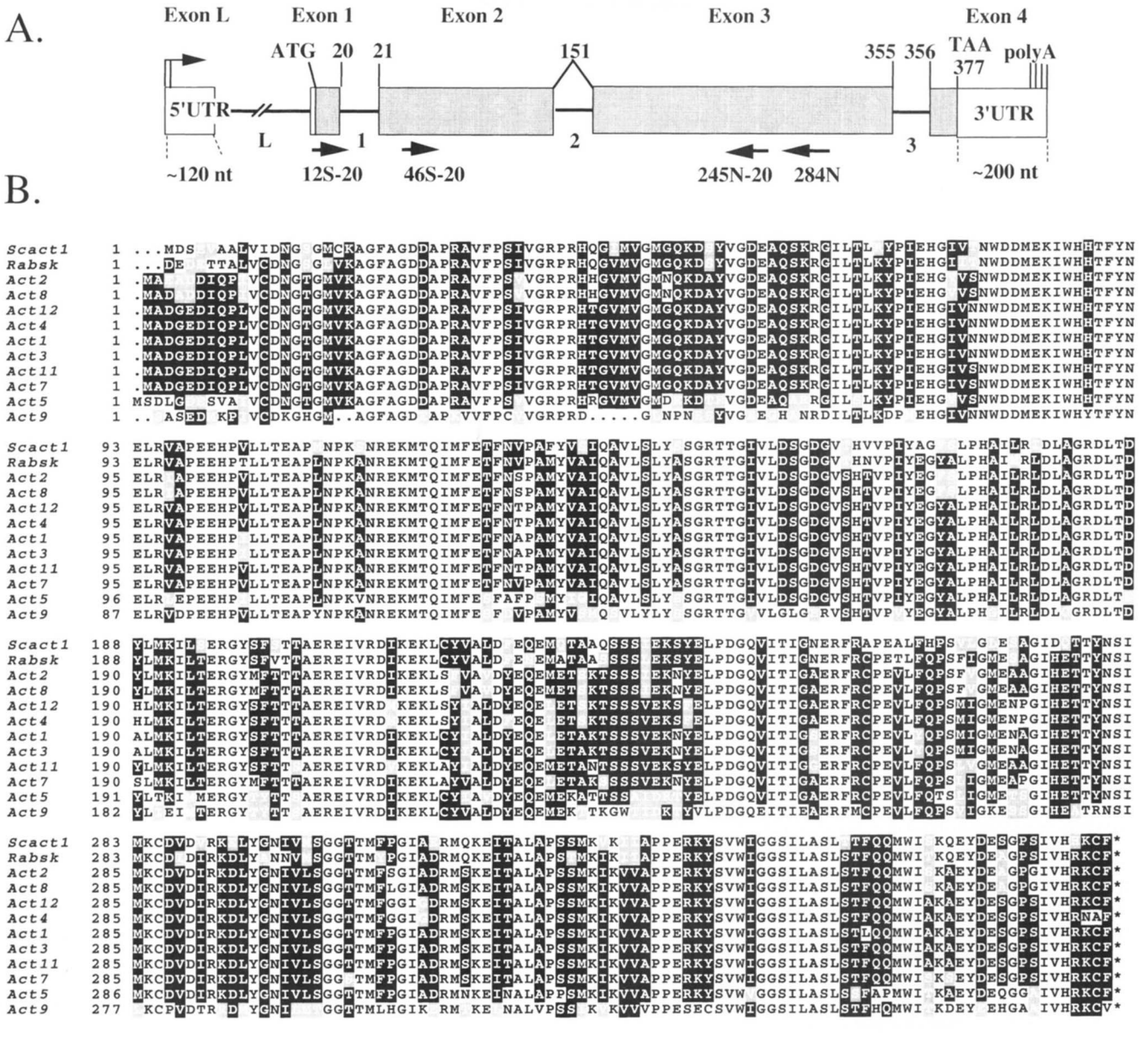

FIGURE 1.—The general structure of the *A. thaliana* actin genes. (A) Physical map of conserved Arabidopsis actin gene organization and location of degenerate actin primers. The locations of the introns are conserved among every gene except for *ACT2*, which does not contain intron 2. Boxes indicate exons. The translated regions are shaded. The numbered arrows below the gene indicate the location of the degenerate PCR primers described in MATERIALS AND METHODS. (B) Multiple alignment of amino acid sequences which were determined from the 10 Arabidopsis actin genes. The sequences of the *S. cerevisiae ACT1* (Sc*ACT1*) and the rabbit skeletal muscle isoform (Rabsk) are included for comparison. Nonconservative substitutions are shaded, and all other substitutions are highlighted in black.

remove excess primer. The product (one-tenth to one-half) was then reamplified in a second reaction with an internal set of primers, 46S-20 and 245N-20. For cDNA clones, two vector primers were used in the first amplification, followed by 46S-20 and 284N in the second amplification. Each degenerate primer (50 pmol) was utilized in 100 μl reactions (SAIKI *et al.* 1988) with a reaction profile of 1′ at 94°, 1′ at 42°, and 1.5′ at 72° for 30–40 cycles.

λ exonuclease digestion and direct sequencing of actin PCR products: We prepared single-stranded DNA from each PCR product according to the lambda exonuclease method, as described in HIGUCHI and OCHMAN (1989). To increase the efficiency of the digestion, excess primer was removed by isopropanol precipitation before digestion, 2 units of lambda exonuclease were utilized per digestion, and the digestion time was increased to 60 min. The 46S-20 primer was then utilized to obtain ~200 bp of diagnostic sequence from the single-stranded digestion product. A total of 87 actin genomic clones and 62 actin cDNAs were characterized with this approach. Unique clones were purified and actin genes and flanking sequences were subcloned into pBluescriptSK+ (Stratagene).

Sequence analysis of genomic clones: Double-stranded plasmid templates were prepared for sequencing by alkaline denaturation (MANIATIS *et al.* 1989). The sequence of the coding region in each gene was obtained with a set of degenerate actin sequencing primers spanning the actin coding sequence in sense and antisense orientations. These primers were 17–23 nucleotides in length and they were positioned with the 5′ (for sense primers) or the 3′ (for antisense prim-

ers) end at codons Met 1, Cys 12, Met 46, Trp 88, Met 125, Val 165, Pro 245, Met 284, Val 339, Thr353, and Ile 369. The following modifications of standard primer annealling protocols were necessary to prevent annealling of degenerate primers to spurious sites: 25 pmol of sequencing primer was annealed to the template with a 30′ incubation at $T_m-10°$ in 6× SSC, followed by a 2′ incubation at $T_m-2°$. It was occasionally necessary to vary the annealling temperatures to obtain optimum results. To remove excess primer, the primer-template hybrid was placed on ice immediately after the $T_m-2°$ shock, precipitated with isopropanol as above, and resuspended in 10 μl 1× Sequenase buffer (United States Biochemical).

The sequence of flanking regions was obtained with unique sequencing primers. All sequencing reactions were performed with the Sequenase 2.0 kit (United States Biochemical) and ^{32}P or ^{35}S adATP (New England Nuclear) according to the supplied instructions. The sequence data was managed with the Gel program in the Intelligenetics suite.

Gel blot analysis of genomic DNA: Arabidopsis genomic DNA was isolated as described in McLean et al. (1988). The DNA was cleaved with restriction enzymes overnight and separated on 0.8% agarose gels at 45 V for 16 hr. DNA was transferred to nylon membranes (Biotrans Plus, ICN) with a vacuum blotter (Hoefer, Inc.) according to the manufacturer's instructions. DNA probes were labeled by the random primer method (Feinberg and Vogelstein 1983) to a specific activity of $0.5-1.0 \times 10^9$ cpm/mg. Low stringency hybridizations were carried out in 6× SSC, 2.5× Denhardt's, 25 mM NaPO4, pH6.5, 0.5% SDS, 35% Formamide for 48 hr at 40°. The blots were washed for 4 × 10 min in 3× SSC, 0.2% SDS, at 50°. Gene-specific hybridizations were carried out in 6× SSC, 5× Denhardt's, 50 mM NaPO4, pH 6.5, 1% SDS, 45% formamide for 48 hr at 56°. Washes were for 4 × 10 min in 0.5× SSC, 0.2% SDS, at 56°. The blots were exposed to X-Omat X-Ray film (Kodak) for 2–4 days with one intensifying screen. Before reprobing, blots were stripped in 0.4 M NaOH at 45° for 45 min, rinsed in 2× SSC, 0.2% SDS, and exposed to X-Ray film for 48 hr to confirm that they were completely stripped.

Alignment of DNA sequences and calculation of nucleotide substitutions: The coding sequences of 22 actin genes (Table 1) were obtained from the Genbank and EMBL databases and aligned in the Genalign program from the Intelligenetics suite. The abbreviated names of these sequences were taken from the original publications. Most of the actin sequences were colinear (377 codons). The following changes were made to correct the exceptions: Codons 26 and 61 were deleted from *ACT7* ORYSA, codon 4 was deleted from *ACT5* ATHAL, codons 130 and 131 were deleted from *ACT2* ORYSA, and codon 332 was deleted from *ACT1* DAUCA. Marker codons (XXX) were inserted at the following gaps in the alignment: Codon 2 in *ACT1* SCER, ACTB RAT, ACTB CHICK, and *ACT8* DICDI; codon 4 in *ACT1* SCER, ACTB RAT, ACTB CHICK, *ACT1* ZMAYS, and *ACT7* ORYSA; codons 4 and 23 in *ACT1* FUC; codons 6 and 34–38 in *ACT9* ATHAL; codon 85 in *ACT1* MAYS and *ACT1* DAUCA; codon 261 in *ACT1* ORYSA; codon 29 in *ACT1* PUISSANT; and codons 330 and 331 in *ACT7* ORYSA.

The divergence between genes was estimated according to the method of Li *et al.* (1986). This method corrects for multiple substitutions and unequal rates of transitions and transversions. It also corrects synonymous and nonsynonymous substitution rates separately.

Gene tree reconstruction: Maximum oarsimony analyses were conducted with PAUP, version 3.1 (Swofford 1990). Third codon positions were deleted from all analyses. All most parsimonious trees were retained in each search. For analyses of ≤11 taxa, branch and bound methods that evaluate all possible trees, but always find the shortest tree, were used. Analyses of ≥20 taxa were done with heuristic search methods with the most exhaustive options available. In all heuristic searches, starting trees were constructed with "simple" addition sequences, followed by branch swapping with tree bisection and recombination. Fifty to 100 random addition sequences were evaluated in each search to avoid "local optima," which could be an artifact of the initial addition sequence. The most parsimonious trees from these searches were then subjected to additional rounds of branch swapping by subtree pruning/regrafting and nearest neighbor inter-

TABLE 1

Actin genes examined

Classification	Gene name	Accession number
A. Plants		
Conferae		
Pinus contorta	*ACT1* PINCO	M36171
Pocaceae		
Zea mays	*ACT1* ZMAYS	J01238
Oryza satvia	*ACT1* ORYSA	Y00279
	ACT2 ORYSA	X15864
	ACT3 ORYSA	X15862
	ACT7 ORYSA	X15863
Solanaceae		
Solanum tuberosum	*ACT58* SOLTU	X55749
	ACT71 SOLTU	X55750
	ACT75 SOLTU	X55753
	ACT97 SOLTU	X55751
	ACT101 SOLTU	X55752
Nicotiana tabacum	*ACT25* NICTOB	X63603
Fabaceae		
Glycine max	*ACT1* SOYBN	J01298
	ACT3 SOYBN	J01297
	ACT4 SOYBN	N/A
Pisum satvium	*ACT1* PISAT	X67666
Apiaceae		
Daucus carota	*ACT1* DAUCA	X17526
Brassicaceae		
Arabidopsis thaliana	*ACT1* ATHAL	M20016
	ACT2 ATHAL	
	ACT3 ATHAL	
	ACT4 ATHAL	
	ACT5 ATHAL	
	ACT7 ATHAL	
	ACT8 ATHAL	
	ACT11 ATHAL	
	ACT12 ATHAL	
B. Animals		
Gallus gallus	*ACTA* CHICK	V01507
	ACTB CHICK	L08165
Rattus rattus	*ACTA* RAT	X06801
	ACTB RAT	V01217
C. Protocists		
Volvox carterii	*ACT1* VOLCA	M33963
Fucus disticus	*ACT1* FUCUS	FDU11697
Dictyostelium discoideum	*ACT8* DICDI	X03284
D. Fungi		
Saccaromyces cerevisiae	*ACT1* SCER	L00026

changes. Neighbor joining and maximum likelihood trees were reconstructed in the PHYLIP package (version 3.5) (FELSENSTEIN 1993). Third codon positions were assigned a weight of zero in every case. The inputs for the neighbor joining program were distance matrices constructed by KIMURA's two parameter method. Other methods of calculating genetic distance were also employed and yielded trees with very similar branching orders.

Construction of space-filling models for plant and human actins: Sequences for human Act-b, Act-g, Act-c, Act-s, Act-a, and Act-h were obtained from GenBank and aligned to identify variable residues. Structural coordinates for the model of rabbit skeletal actin-DNaseI complex (KABSCH *et al.* 1990) were retrieved from the Bookhaven Database. Rabbit skeletal actin is identical in sequence and presumably in structure to human Act-s and was used as the basis for building a model plant actin. The DNaseI molecule was deleted from this model using Nitro (Tripos Assoc. Inc.). Building plant *ACT2* model structure required making 52 amino acid substitutions within the above skeletal muscle actin. Changing residue 7 from ala7 in Act-s to pro7 in *ACT2* could not be done without making a minor change in conformation of the peptide backbone. Residues which varied within the plant or animal families were located within the Act-s and *ACT2* structures, respectively, and displayed in a spacefilling model.

	ACT1	*ACT2*	*ACT3*	*ACT4*	*ACT5*	*ACT7*	*ACT8*	*ACT9*	*ACT11*	*ACT12*	*SCER*
ACT1	0	0.0666	0.0023	0.0251	0.0994	0.0420	0.0647	0.1787	0.0317	0.0257	0.1550
ACT2	1.297	0	0.0660	0.0635	0.1018	0.0426	0.0046	0.1814	0.0518	0.0648	0.1619
ACT3	0.6055	1.133	0	0.0269	0.0994	0.0420	0.0641	0.1800	0.0347	0.0251	0.1561
ACT4	1.242	1.677	1.432	0	0.1026	0.0494	0.0604	0.1789	0.0293	0.0046	0.1570
ACT5	1.452	1.969	1.366	1.444	0	0.1068	0.0973	0.1608	0.0935	0.1059	0.1887
ACT7	0.9957	1.176	1.192	1.483	1.776	0	0.0408	0.1806	0.0403	0.0469	0.1616
ACT8	1.067	0.5571	1.228	1.351	1.603	1.045	0	0.1769	0.0493	0.0598	0.1634
ACT9	1.552	1.562	1.458	1.434	0.3960	1.676	1.292	0	0.1729	0.1759	0.2499
ACT11	1.100	1.097	1.289	1.438	1.342	1.250	1.218	1.170	0	0.0293	0.1571
ACT12	1.447	1.722	1.331	0.8351	1.535	1.569	1.764	1.667	1.426	0	0.1641
SCER	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	0

FIGURE 2.—Nonsynonymous and synonymous nucleotide substitutions among Arabidopsis actin genes. Nonsynonymous (RNS) and synonymous (SNS) substitutions (above and below the diagonal, respectively) were calculated for all pairwise comparisons of the Arabidopsis actin protein coding sequences (see MATERIALS AND METHODS). The numbers shown were corrected for multiple hit kinetics and are given as a fraction of the total number possible substitutions (*i.e.*, in the RNS comparison of *ACT1 vs. ACT2* the value of 0.0666 means that 6.66% of the total possible changes occured). Values which corrected to numbers >2.0 (200%) are listed as NAN.

RESULTS

Isolation of 10 distinct Arabidopsis actin sequences: By screening several Arabidopsis libraries with heterologous actin probes, we isolated 10 distinct Arabidopsis actin genes, including the previously isolated *ACT1* gene. Genomic DNA clones containing *ACT1, 2, 3, 4, 5, 7, 8, 9, 11,* and *12,* and cDNA clones of *ACT2, 7,* and *8* were isolated. We determined the complete nucleotide sequence of each gene, including the coding sequence, introns, ~1200 bp of 5′ flanking sequence, and ~400 bp of 3′ flanking sequence. We also isolated putative alleles of *ACT2* and *ACT4* (*ACT2A* and *ACT4B,* respectively). *ACT2A* contains a 1-bp, silent substitution that changes codon 104 from CCT to CCC, and *ACT4B* contains a 1-bp deletion in codon 120 that creates a stop codon in the open reading frame. No other substitutions were detected in these genes. During the screening, we isolated each of the 10 genomic clones at least four times. This suggested that we had screened the libraries to the point of significant redundancy and that no additional actin genes would be found.

Comparison of Arabidopsis actin gene organization: All but one of the Arabidopsis actin genes contained three small introns at identical locations in the coding sequence (Figure 1A). Intron 1 separates codons 20 and 21, intron 2 splits codon 151, and intron 3 is located between codons 356 and 357. Every previously reported plant actin gene contains three coding sequence introns at these exact locations. In *ACT2,* however, intron 1 appears to have been precisely removed. Introns 2 and 3 are present in the conserved locations in *ACT2,* and the *ACT2* protein coding sequence is intact and completely colinear with the other Arabidopsis actin genes.

Each of the Arabidopsis actin messages contains a 3′ untranslated region (UTR) of 180–220 nt and a short, untranslated leader region that is split by a large intron −9 to −12 bp upstream of the ATG initiation codon. Detailed structures of the 5′ and 3′ UTRs of the expressed genes are presented in separate manuscripts (AN *et al.* 1996; HUANG *et al.* 1996; Y.-Q. AN, J. M. MCDOWELL, S. HUANG, E. C. MCKINNEY, S. CHAMBLISS and R. B. MEAGHER, unpublished data; J. MCDOWELL, Y.-Q. AN, E. C. MCKINNEY, S. HUANG and R. B. MEAGHER, unpublished data).

Divergence among the Arabidopsis actin genes: The 10 Arabidopsis genes were easily recognizable as actin, differing by only 14.9–25.1% RNS (nonsynonymous substitutions) from the *Saccharomyces cerevisiae* actin 1 gene (*SCER*; Figure 2 and Table 2), and by 9.2–23.0% from animal muscle and cytoplasmic actin genes. The 10 Arabidopsis actin genes differed from each other by 0.4–18.5% RNS. However, for the reasons described in the following paragraph, two of the Arabidopsis sequences, *ACT5* and *ACT9,* are probably pseudogenes. When they are dropped from the comparison, the RNS values for the remaining eight Arabidopsis genes are considerably lower (Table 2). Three closely related pairs of Arabidopsis actin genes were found that differed by only one conservative substitution in their derived amino acid sequences (Figure 1): *ACT2* and *ACT8* (Glu to Asp substitution at residue 3), *ACT1* and *ACT3* (Leu to Phe substitution at position 354), and *ACT4* and *ACT12* (Tyr to Phe at position 225) (Figure 1). These three pairs were also closely related in terms of synonymous substitutions (SNS), differing by 55.7, 60.6, and 83.5% SNS, respectively.

Most of the other Arabidopsis actin genes differed from each other by 2.6–6.8% RNS and >100% SNS. *ACT5* and *ACT9* were exceptionally divergent. *ACT5*

contained a one codon insertion at the 5′ end and was 9.4–16.1% diverged in RNS from the other eight Arabidopsis actins. *ACT9* differs by 16.1–18.1% RNS from the other Arabidopsis actins. The first in-frame ATG in the *ACT9* coding sequence occurs 18 bp downstream of the translation initiation site, which is conserved among the other nine Arabidopsis actin genes. This deleted the first six conserved actin codons from the *ACT9* protein. *ACT9* also contains a 5-codon deletion of conserved actin codons 43–47. *ACT5* and *ACT9* were isolated on the same lambda clone and were more similar to each other (16.1% RNS and 39.6% SNS) than to the other Arabidopsis actin genes. Interestingly, we were unable to detect mRNA from either of these genes on RNA gel blots or with reverse-transcriptase-mediated PCR (data not shown). The lack of detectable expression from *ACT5* and *ACT9,* along with their unusual level of RNS divergence, suggested that both genes were nonfunctional.

Pairwise alignments of the introns and flanking regions of the 10 genes (data not shown) revealed no significant similarity, except for several small blocks of identity near the transcription start sites among the most closely related gene pairs. This indicated that gene-specific probes could be constructed from both 5′ and 3′ flanking regions.

Genomic structure of the Arabidopsis actin gene family: To estimate the minimum complexity of the actin gene family in Arabidopsis, "universal" actin coding sequence probes from the Arabidopsis *ACT2* gene and the maize actin gene, *MAC1* (SHAH *et al.* 1983), were hybridized under low stringency to Southern blots of Arabidopsis genomic DNA. The high degree of conservation among all actin genes indicated that these probes would detect all of the Arabidopsis actin genes (HIGHTOWER and MEAGHER 1985). The *ACT2* probe hybridized to 10–14 bands in each digest (Figure 3). Similar results were obtained with the maize actin probe (data not shown).

To determine the copy number of each gene and to assign identities to the bands that hybridized to the universal actin probes, we made specific probes for each of the 10 genes and hybridized them to identical genomic DNA blots under moderate stringency. The lack of

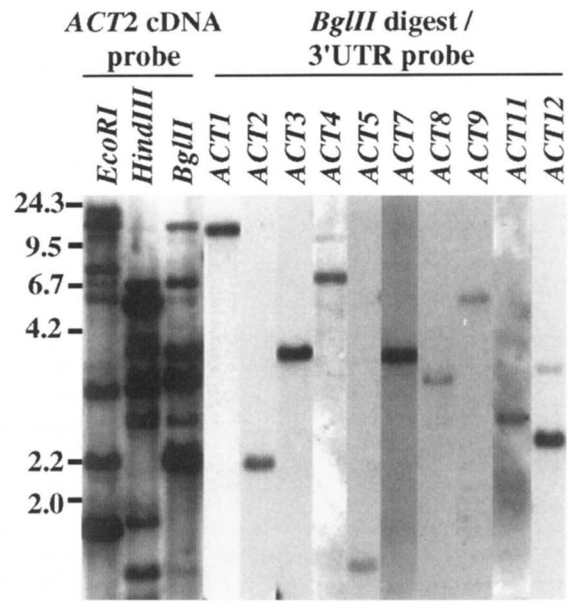

FIGURE 3.—Southern blot analysis of Arabidopsis genomic DNA suggested that each gene was unique in the genome. Arabidopsis genomic DNA (5 μm) was digested with the restriction endonucleases indicated, separated on a 0.8% agarose gel, and transferred to a nylon membrane. The blot shown in lanes 1–3 was hybridized to an *ACT2* cDNA insert at low stringency. Lanes 4–13 show *Bgl*II digests of Arabidopsis DNA, which were hybridized to gene-specific 3′ flanking sequence probes for each of the 10 actin genes, under moderate stringency.

flanking sequence similarity among the Arabidopsis actin genes suggested that probes from this region should be gene specific. This was tested by hybridization to dot blots containing DNA from each Arabidopsis actin gene (data not shown); no cross-hybridization was observed. The probes were then hybridized to DNA blots that were replicas of the ones used for the universal actin hybridizations. Each gene-specific probe hybridized to a single band, again confirming the gene-specificity of the probes and suggesting that there were no closely related copies of any of the 10 genes in Arabidopsis.

We compared the banding patterns from the gene-specific hybridizations with those obtained with the "universal" Arabidopsis and maize actin coding sequence probes. Each of the bands detected in the gene-

TABLE 2

Summary of RNS values for the eight functional Arabidopsis actin genes

Comparison	RNS[a]	Comparison	RNS
Plant *vs.* plant[b]	0.4–12.3	Plant *vs.* animal	9.2–15.3
Arabidopsis *vs.* Arabidopsis	0.4–6.8	Plant *vs.* protist[c]	9.8–14.9
Animal *vs.* animal	0.1–6.8	Plant *vs.* fungal[d]	14.3–18.7

[a] Replacement nucleotide substitutions (RNS), calculated among all pairwise comparisons according the method of LI, WI, and LUO. Values are given in frequence of substitution per 100 nucleotides. They are corrected for multiple hit kinetics.

[b] *ACT5* and *ACT9* were eliminated from the comparisons because they are putative pseudogenes.

[c] Dictyostelium, Volvox, and Fucus.

[d] *S. cerevisiae.*

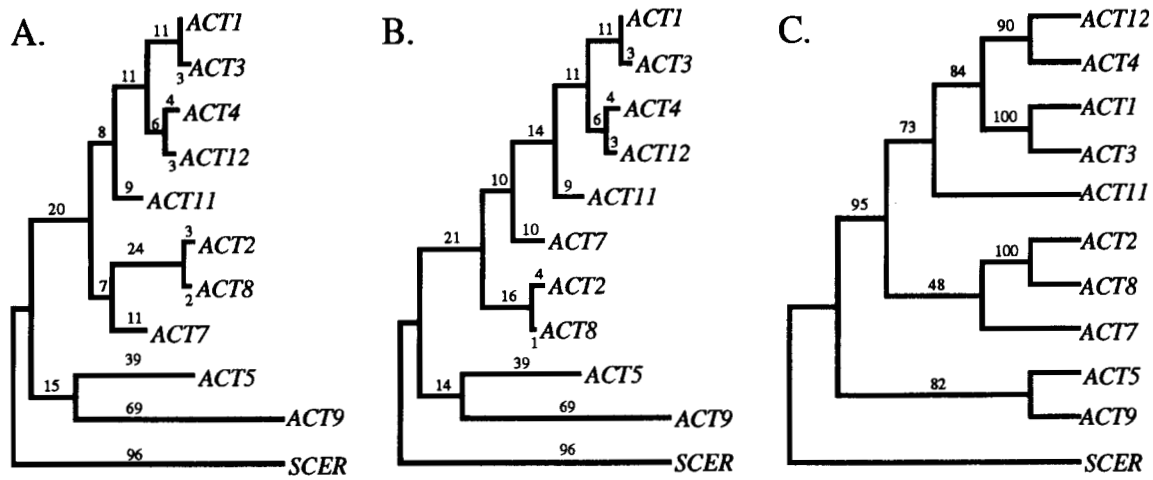

FIGURE 4.—Evolutionary relationships of the Arabidopsis actin genes. Gene trees A and B are maximum parsimony trees that were generated in PAUP using branch and bound analysis. Gene tree C is a branch-and-bound, bootstrap consensus tree constructed in PAUP, with 100 bootstrap replicates. Branch lengths in trees A and B are scaled according to the number of substitutions that were estimated to have occurred along each branch. For an approximate scale to these distances *ACT1* differes from *ACT2* by 6.7% RNS (Figure 2).

specific hybridizations corresponded to one of the bands that hybridized to the conserved coding sequence probes at low stringency, suggesting that most or all of the Arabidopsis actin genes detectable by standard hybridization techniques had been isolated. The *Bgl*II fragments are presented in Figure 3 as an example of the results we observed. Figure 3 demonstrates that all of the bands that were detected by the general actin probes were accounted for in the gene-specific hybridizations. We were also able to account for all of the bands in the *Xba*I and *Bgl*II digests with this approach (data not shown). However, a single band in the *Eco*RI (~6.4 kb) and *Hin*cII (~5.0 kb, data not shown) digests did not hybridize to 5′ or 3′ probes from any of the 10 genes, even though they hybridized to the general actin coding sequence probe.

Phylogenetic relationships of the Arabidopsis actin genes: The evolutionary relationships of the Arabidopsis actin genes was determined by several different tree-building methods. Maximum parsimony and compatibility algorithms both produced two minimum length trees that were essentially identical. The maximum parsimony trees from PAUP are shown in Figure 4, A and B, and referred to as tree A and tree B, respectively. Tree C is a majority rule consensus tree from 100 bootstrap replicates in PAUP. The four similar pairs of genes, *ACT1/3, ACT4/12, ACT2/8,* and *ACT5/9,* form distinct clades in all trees. *ACT11* occupies a distinct branch on the tree that is basal to the *ACT1/3* and *4/12* clades. Trees A and B differ with respect to the placement of *ACT7.* In tree A, *ACT7* is a sister group to the *ACT2/8* clade. In tree B, *ACT7* is basal to the clade containing *ACT1, 3, 4, 11,* and *12.* Maximum parsimony analysis (DNA-PARS, PHYLIP 3.5), and compatibility analysis (DNA-COMP, PHYLIP 3.5) produced two most parsimonious trees that were identical to trees A and B. The topology of trees produced by the neighbor joining method and the maximum likelihood method were identical to that of tree A. UPGMA produced a tree identical in branching order to tree B. The *S. cerevisiae* outgroup roots the trees between the *ACT5/9* clade and the branch leading to the rest of the actins. Because *ACT5* and *9* may be evolving without constraint on RNS and at a different rate from the other sequences, we reconstructed trees that did not contain either of these taxa. Every tree building algorithm we used produced one optimal tree that was identical to tree A and was rooted between the *ACT2/7/8* clade and the remainder of the taxa, suggesting that *ACT7* is a sister group to the *ACT2/8* clade.

The relationship of the Arabidopsis genes to other eukaryotic actin genes, including all of the available plant actin sequences, is shown in Figure 5. Genes from monocots (rice and maize), dicots (Arabidopsis, soybean, pea, tobacco, potato, and carrot) and one gymnosperm (lodgepole pine) were included in the analysis, as well as representative actin sequences from animals, protists, and fungi. The complete coding sequence of each gene was analyzed, except for the pine actin gene, which contained 471 nt from the 3′ end (41% of the coding region). Third codon positions were disregarded in all analyses because most were randomized among sequences that had been separated for hundreds of millions of years, and would thus contribute an unacceptable level of homoplasy to the data set. The distribution of 1000 random trees produced from this data set in PAUP was signigicantly skewed to the right ($g1 = -0.8537$). This indicated that a strong phylogenetic signal was present in the data set.

Figure 5, A and B, shows phylograms that were produced by PAUP and the neighbor joining method. Bootstrap values ≥70% are indicated in each tree. The plant actin sequences in both trees are monophyletic with respect to nonplant actins. The deduced ancestral

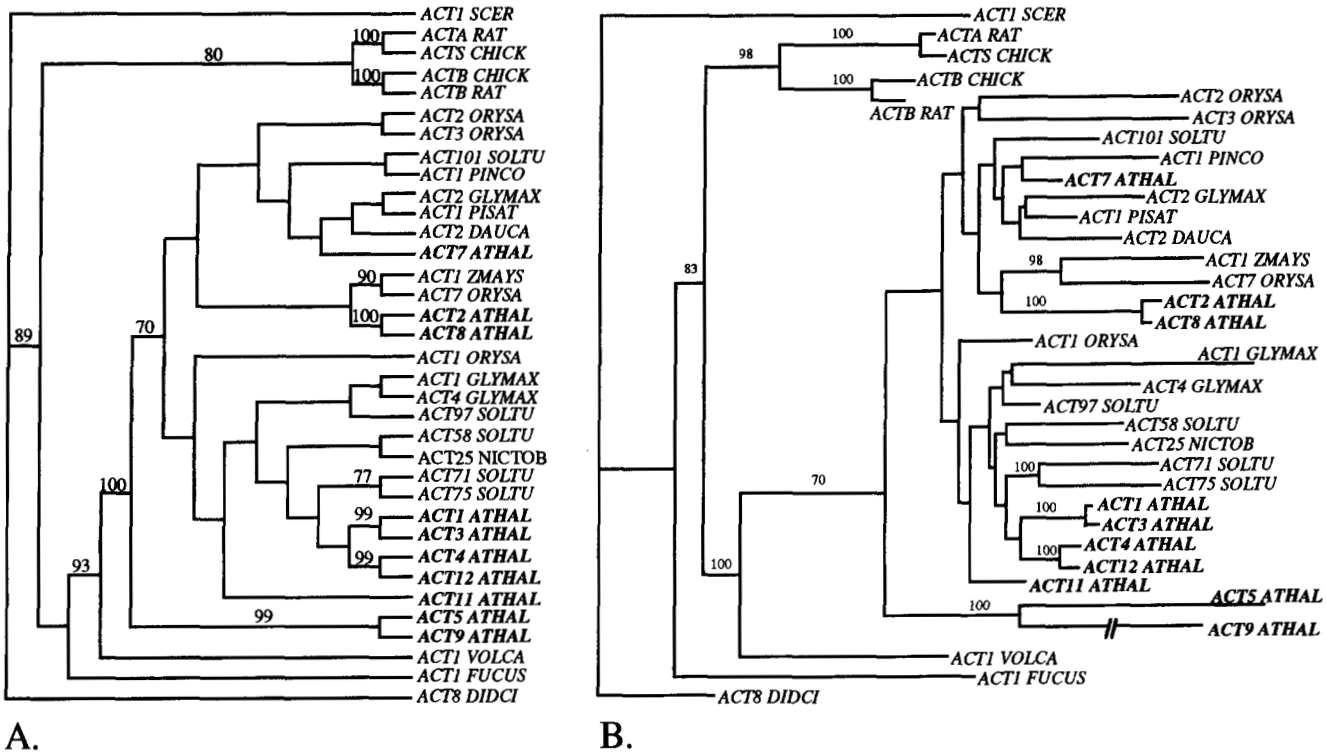

FIGURE 5.—Phylogenetic relationships of the Arabidopsis actin genes to actin genes from animals, protists, fungi, and other plants. The gene trees shown are based on majority-rule consensus from 100 bootstrap replicates. Tree A resulted from heuristic searching in PAUP 3.1. Tree B is a neighbor joining tree from distance matrices that were constructed according to KIMURA's two-parameter method. First and second codon positions from the entire protein coding region of each gene were used in each reconstruction. Bootstrap values >70% are shown. Branch lengths are scaled in tree B as estimated by the neighbor joining program and are proportional to genetic distance. For an approximate scale to these distances *ACT1* differes from *ACT2* by 6.7% RNS and from yeast actin, SCER, by 15.5% (Figure 2). Branch lengths are not scaled in tree A.

sequence of plant actins is most closely related to the actin gene from the green alga, *Volvox carterii*. The plant genes are separated into two major clades (*i.e.*, classes, see DISCUSSION) that contain sequences from both monocots and dicots. Each of these two clades contains subclades composed of genes from distantly related species. The six Arabidopsis actin genes subclasses are dispersed throughout the tree. Five of the six Arabidopsis subclasses are more closely related to genes from other plant species than they are to each other, suggesting that they originated from lineages that split off before those species diverged. The *ACT2/8* clade, for example, groups with genes from maize and rice. *ACT7* groups with genes from pea, carrot, potato, and pine. *ACT11* is the basal member of a clade that contains genes from soybean, potato, tobacco, and Arabidopsis. This clade also forms a sister group with a rice actin gene. *ACT1/3* and *ACT2/4* are the only two Arabidopsis actin clades that are more closely related to each other than to genes from other species. The *ACT5/9* putative pseudogenes occupy a basal position relative to all of the plant genes.

The relationship of the Arabidopsis sequences to the other plant genes, as proposed by trees in Figure 5, suggests that the Arabidopsis sequences, and indeed all the plant sequences, originated from a series of ancient gene divergences. However, bootstrap support for many of these larger groups generated by these gene duplications, particularly the one that separated the two major groups of plant and Arabidopsis actin genes, was weak. We used a second approach to evaluate two primary competing hypotheses: that the Arabidopsis actin genes had an ancient origin leading to bifurcations in the tree or diverged relatively recently and thus are monophyletic. The hypotheses were evaluated by "forcing" the Arabidopsis members to remain monophyletic (*i.e.*, to form a lineage that included no other plant actin genes) during tree reconstruction in PAUP, via the "enforce monophyly" option. We then compared the length, as determined by PAUP, of the most parsimonious trees satisfying this constraint with the length of the tree in Figure 5A. We thus identified the six most parsimonious in which the Arabidopsis actin genes formed a separate lineage (length = 831 steps). These trees were longer than the original tree (length = 814 steps). We then repeated the procedure and forced each plant actin gene family to be monophyletic. When this was done for all species simultaneously, we found 11 most parsimonious trees that conformed to this constraint. These trees were 73 steps longer than the one made under the original hypotheses. Based on the criterion of maximum parsimony, these results strongly favor the polyphylogenetic hypothesis suggested in Figure 5A.

Survey of concerted evolution among the Arabidopsis

actin genes: We observed no long stretches of nucleotide identity or near identity among the Arabidopsis actin genes, suggesting that no recent, large scale gene conversions had occurred. This observation was confirmed statistically with a G-test (as described in DROUIN and DOVER 1990), which revealed no evidence of gene conversion (data not shown). We then employed a phylogenetic approach, based on compatibility analysis (DROUIN and DOVER 1990), to search for small scale nonreciprocal exchanges. The phylogeny of the Arabidopsis actin sequences (excluding *ACT9*) was estimated using the DNA-COMP program (PHYLIP 3.2). Third position sites were again ignored in the phylogeny reconstruction. Each nucleotide position, including third codon positions, was then examined for substitutions inconsistent with the hypothesized evolutionary relationships from the trees. The rationale for this approach was that a cluster of contiguous substitutions that were incompatible with the phylogeny (*i.e.*, could not be explained by shared common ancestry), and shared between a pair of sequences, would indicate that a conversion event has occurred. As noted by DROUIN and DOVER, this method could potentially enable resolution of very small (*i.e.*, <10 base pairs) gene conversions.

Two optimal trees were produced that were identical to the trees in Figure 4. The location of sites incompatible with each phylogeny are shown on a multiple alignment of the nine actin coding sequences shown in Figure 6. Substitutions at 265 of the 1137 sites were incompatible with either or both phylogenies. In general, the incompatible sites were dispersed over the coding sequence, and the majority of incompatible sites were in third codon positions. No clusters of five or more consecutive, incompatible sites were observed. Only one cluster of four incompatible sites was observed, at positions 657–660. The incompatible patterns of sharing at each site in this cluster involved a different pair of genes: *ACT4/12* and *ACT2* at site 657, followed by *ACT7* and *ACT11* at site 658, then *ACT1/3* and *ACT5* at site 659, and finally *ACT2/8* and *ACT1/3* at nucleotide 660. A cluster of seven incompatible sites over 10 nucleotides was located at positions 201–210; the only pair of sequences that shared more than two substitutions in this region were *ACT4* and *ACT12*, which are directly related by descent. Ten regions of three consecutive, or three out of four, incompatible sites were also observed. No sharing of contiguous sites or groups of contiguous sites within these regions that could not be explained by descent from a common ancestor were observed in any of these regions. Thus, nonreciprocal recombination does not appear to have significantly influenced Arabidopsis actin gene evolution.

Structural variability among the Arabidopsis actins: One part of our working hypothesis predicts that ancient actin subclasses might encode functionally divergent proteins. In a previous study, MCLEAN and MEAGHER (MCLEAN *et al.* 1990) noted a striking degree of variability among three soybean actin genes at positions encoding charged amino acids. The level of amino acid divergence within the Arabidopsis actin protein family was compared to that among six members of the human actin protein family where minor functional differences have been noted among the proteins. The Arabidopsis *ACT5* and *ACT9* sequences were eliminated from the analysis because they may encode disfunctional proteins no longer under selection. All substitutions involving lys, arg, his, asp, and glu (charged), asn and gln (strongly polar), or proline (an imino acid that dramatically changes the peptide backbone) were identified. There were eight amino acid positions with these substitutions among the six human actin isoforms. These were mapped on the human Act-s protein and are shown in Figure 7, A–D. Two to four substitutions were observed between any one pair of human actins. Most of the variation was confined to the amino terminal end of the protein, and all of the substitutions were very conservative asn/gln or asp/glu interchanges. The five Arabidopsis actin subclasses exhibited charged residue or proline residue variation at 11 positions and these are mapped on a model of the Arabidopsis ACT2 protein (Figure 7, E–H). There were three to eight differences between any two Arabidopsis actin subclasses. Only two of these differences could be classified as conservative (*i.e.*, asp at the third position to glu, asp3/glu, and glu at the forth position to asp, glu4/asp interchanges). The other nine changes all had potentially significant affects on the character of the protein (*e.g.*, his41/thr, asn48/gly, ser77/asn, tyr188/his/ala/ser, lys231/asn, asn239/ser, ala270/asn, ala272/pro, ser307/leu/gly/pro). In further contrast to the animal actins, these substitutions were distributed widely over all surfaces of the plant actin molecule. Interestingly, many of these charged residue substitutions were conserved within a phylogenetic class (*e.g.*, *ACT2, 8,* and *7, vs. ACT1, 3, 4, 12,* and *11*) or closely related subclasses of plant Arabidopsis actins (see DISCUSSION). While the many of the remaining amino acid substitutions among the Arabidopsis actins were very conservative (*e.g.*, ile/val, leu/met, ala/gly, tyr/phe), there were more substitutions with potentially drastic affects on the plant actin structure (*e.g.*, ser129/thr/ala/val, met199/ser, ser217/cys/ala, ser358/ala) than are exhibited among the animal actins (not shown).

DISCUSSION

Actin is encoded by 10 genes in *A. thaliana*: A primary goal of this study was a thorough analysis of actin gene family structure in Arabidopsis. Hybridization of actin coding sequence probes to Arabidopsis genomic DNA blots revealed a gene family of moderate size. Through an exhaustive screening of several Arabidopsis libraries, we isolated 10 unique actin sequences, including the previously isolated *ACT1* gene, and two putative alleles. Four lines of evidence suggested this collection contained most, and possibly all, of the Arabidopsis actin genes: 1) Among the 87 genomic clones that were char-

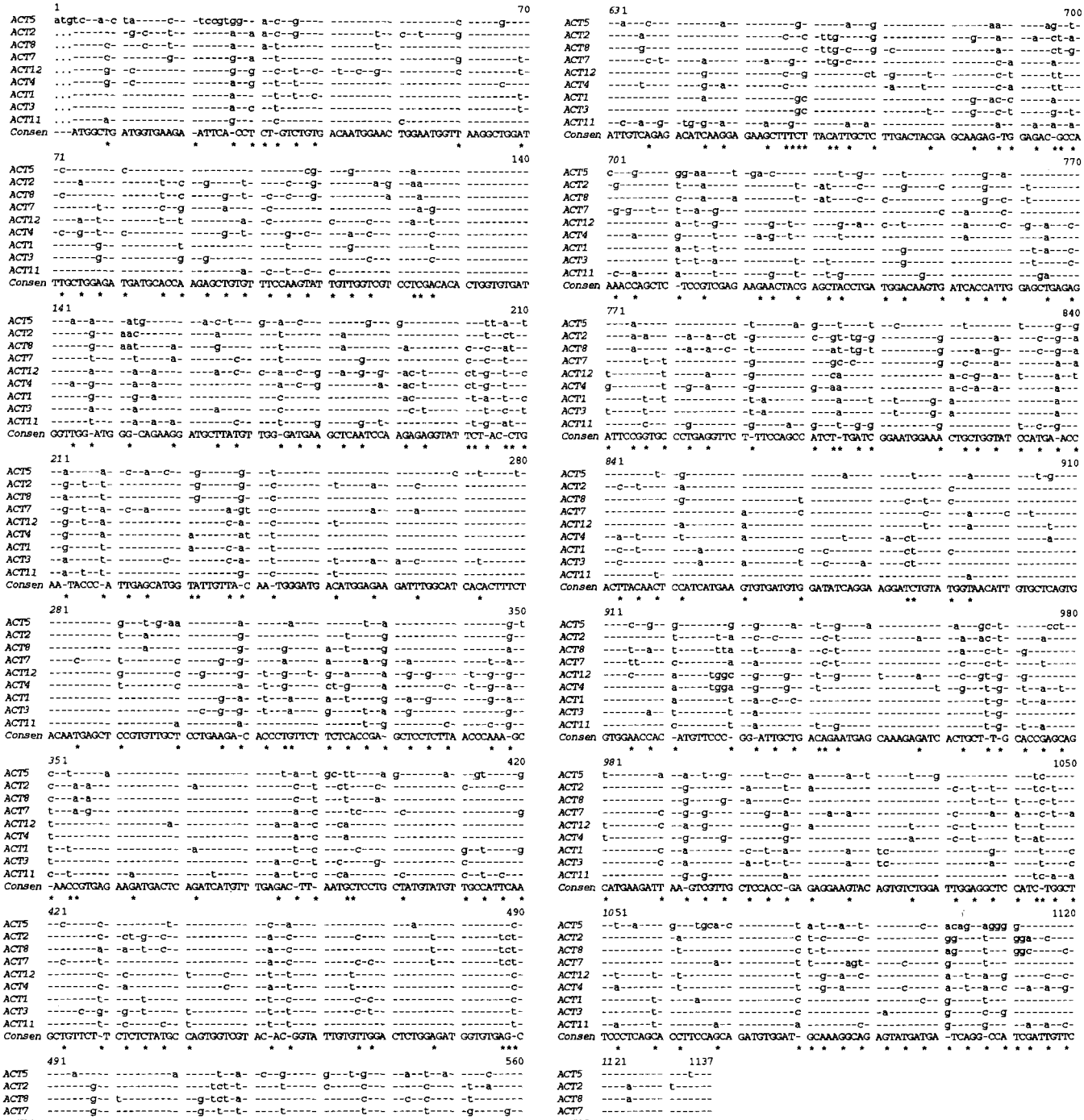

FIGURE 6.—Compatibility analysis suggested that the Arabidopsis actin coding sequences evolved independently and were not subjected to mechanisms of sequence homogenization. Nucleotide substitutions among the Arabidopsis actin genes that are incompatible with their hypothesized phylogenetic relationships are shown on a multiple alignment of the Arabidopsis sequences. The multiple alignment displays a consensus sequence in the bottom row. All nucleotides which are in agreement with the consensus are marked with a dash. Sites that are incompatible with Tree A and/or Tree B in Figure 4 are marked with an asterisk below the consensus sequence.

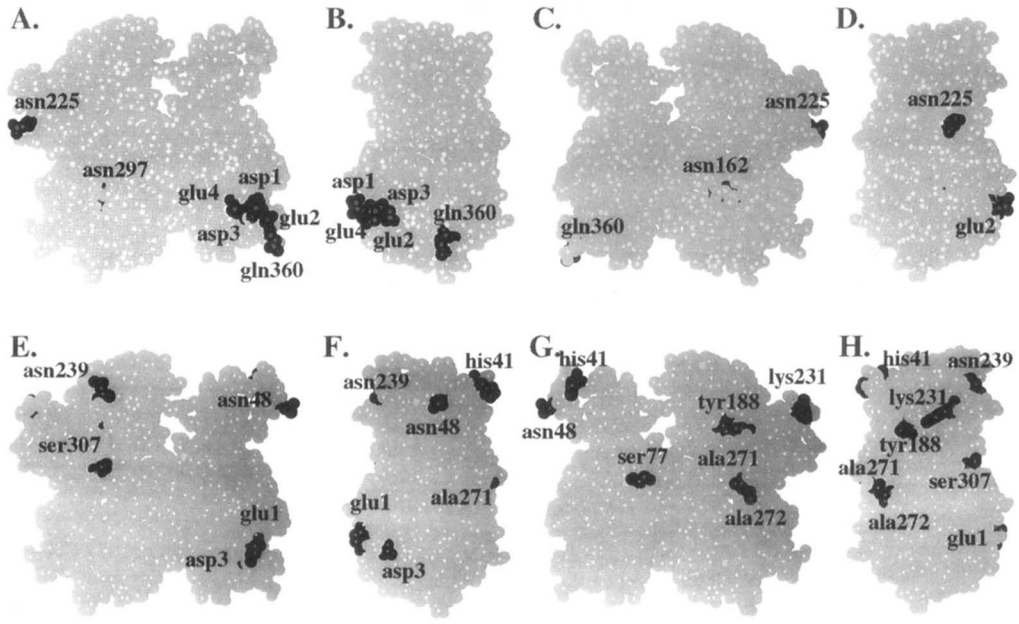

FIGURE 7.—Nonconservative amino acid substitutions within the human and Arabidopsis actin protein families. Substitutions involving charged or strongly polar residues, as well as imino- (pro)/α-amino acid interchanges, were mapped on spacefilling models of human and Arabidopsis actin isoforms. The models were based on the known three-dimensional structure of the rabbit skeletal muscle actin. (A–D) Front, right side, back, left side view, respectively, of the spacefilling model for human skeletal actin. Substitutions among the six human actin isoforms (α, β, γ, c, s, h) are highlighted in black. (E–H) Front, right side, back, and left side view respectively, of a space-filling model for Arabidopsis *ACT2*. The *ACT2* structure was built based on the human Act-s structure. Substitutions among the eight expressed Arabidopsis actins (*ACT1, 2, 3, 4, 5, 7, 8, 11, 12*) are shown in black. Substitution of Arabidopsis pro7 for the human ala7 caused significant structural changes in the actin backbone in that region that were not verified by any independent method.

acterized, each of the 10 genes was isolated at least four times. This indicated that our screening had reached the point of extreme redundancy and that no more Arabidopsis actin genes remained to be isolated from the libraries we examined. 2) The hybridizations with flanking sequence probes indicated that each of the 10 genes was unique, with no closely related homologues. 3) Hybridizations with actin coding sequences at low stringency were sensitive enough to detect very dissimilar actin genes on Southern blots and during library screening. This was evidenced by our ability to detect the highly divergent *ACT9* gene, which shares only 70% nucleotide identity with the coding sequence probes. 4) We were able to use flanking sequence probes to assign identities to almost every one of the bands from five different restriction digests that hybridized to the conserved actin coding sequence probes. However, one band in two of the five digests did not hybridize to any of the flanking or coding sequence probes. This fragment may represent a distinct actin gene which has not been identified but is more likely an allelic variant (*e.g., ACT2A* or *ACT4B)* or partial duplication of one of the genes. It is also possible that Arabidopsis contains actin-related genes similar to the genes characterized in other organisms that have 30–55% amino acid identity with actin (CLARK and MEYER 1992; LEES-MILLER *et al.* 1992a,b; SCHWOB and MARTIN 1992). While these actin-related genes would not be expected to hybridize to the actin coding sequence probes used due to their overall level of divergence, it is possible that they might share some block of sequence homology sufficient to give this result. Thus, we cannot conclusively state that all actin-related sequences in Arabidopsis have been identified. However, the available evidence suggests that the 10 genes we have isolated include most, if not all, of the Arabidopsis genes with strong similarity to "conventional" actin genes.

These results establish that the Arabidopsis has the smallest actin gene family of any plant yet examined. The actin gene family in Petunia has been estimated to contain >100 members (BAIRD and MEAGHER 1987) and potato, tobacco, and soybean also appear to contain dozens of actin genes (DROUIN and DOVER 1990; MCELROY *et al.* 1990; THANGAVELU *et al.* 1993). Perhaps Arabidopsis contains the minimum number of actin genes higher plants require. However, the presence of three closely related gene pairs in the Arabidopsis actin gene family suggests that even its relatively "streamlined" structure contains some redundancy. The 10 genes we have characterized fall into six distinct subclasses. This places the Arabidopsis actin gene family at approxiamtely the same level of complexity as the actin gene families from most metazoans (about six functional actin gene subclasses). This observation implies that the remarkable complexity of the actin gene families in other plants simply reflects increased redundancy, rather than a requirement for dozens of uniquely specialized genes.

Intron 1 was probably deleted from the *ACT2* gene: A notable feature of plant actin genes is that their intron-exon organization is conserved at the kingdom level. Every plant actin gene, with the exception of the processed pseudogene PoAc99 (DROUIN and DOVER 1987), contains three introns that interrupt the coding sequence at identical locations (MEAGHER and WILLIAMSON 1994). In contrast, metazoan actin genes contain one to seven introns at 12 different positions (SHETERLINE and SPARROW 1994). We found that all of the Arabidopsis actin genes conformed to the expected intron-exon organization except for *ACT2*, in which intron 1 is missing. The ACT2 reading frame was preserved, and introns 2 and 3 are found at the expected sites.

Because intron 1 is present at the same location in every other known plant actin gene, the most parsimonious hypothesis is that it has been deleted from *ACT2*. The *ACT8* gene, which is closely related to *ACT2*, contains intron 1 at the conserved location. This indicates that the loss of intron 1 from *ACT2* was relatively recent. The date of this event could be estimated by examining the organization of *ACT2* and *ACT8* orthologs in other Brassicacae. *ACT2* is one of the most highly expressed Arabidopsis actin genes (Y.-Q. AN, J. M. MCDOWELL, S. HUANG, E. C. MCKINNEY, S. CHAMBLISS and R. B. MEAGHER, unpublished results), accounting for 50% of the actin mRNA in some vegetative organs. The loss of this intron in a highly expressed gene demonstrates that introns can be deleted without affecting the functional capability of the gene. A mechanism for the precise removal of an intron by RNA-mediated recombination has been described in *S. cerevisiae*, and requires an endogenous retroviral activity (DERR and STRATHERN 1991). Because Arabidopsis retroelements have been identified (VOYTAS and AUSUBEL 1988), it is plausible that intron 1 was deleted by recombination with an *ACT2* cDNA. *ACT2* mRNA is the most abundant actin mRNA in the plant (Y.-Q. AN, J. M. MCDOWELL, S. HUANG, E. C. MCKINNEY, S. CHAMBLISS and R. B. MEAGHER, unpublished data) and is expressed in meristematic cells that subsequently give rise to reproductive tissues and germ cells in the plant.

Lack of evidence for gene conversion during Arabidopsis actin gene evolution: Multigene families can undergo sequence homogenization, often referred to as "concerted evolution." DOVER and others hypothesized that homogenization of related sequences could result from a number of mechanisms, including unequal crossing over, replication slippage, transposition, and gene conversion. Gene conversion, in particular, can significantly affect the evolution of plant gene families (MEAGHER *et al.* 1989). We employed both statistical and phylogenetic tests to look for putative conversions among the Arabidopsis actin genes, and found no evidence for even small nonreciprocal recombinations. The members of each of the most closely related pairs (*ACT1/3*, *2/8*, and *4/12*) did share extensive similarity, which may have arisen through relatively recent conversions. However, the lack of similarity between the genes in flanking regions and introns argues against this hypothesis. In the compatibility analysis, the clusters of sites that were inconsistent with the phylogeny were few, far between, short in length, and composed of sites shared among different pairs of genes. No significant clustering of shared incompatible sites that would have been indicative of even small-scale conversions was observed, with the exception of three regions that contained three dispersed, incompatible sites shared among the same pairs of genes. These three regions may represent small gene conversions; however, the incompatible sites in these regions are all third codon positions, which makes convergent evolution by random drift a reasonable alternative hypothesis. Of more interest is the region between positions 127 and 142, which is completely conserved among all nine genes. The lack of divergence in this region, even at silent sites, suggested that a "global" homogenation of some sort had taken place or that the sequence was conserved at the nucleotide level and not just for its amino acid coding capacity.

In view of the strong influence of gene conversion on the evolution of other plant actin gene families such as RbcS and the ubiquitins (CALLIS *et al.* 1995; MEAGHER *et al.* 1989), the apparent lack of conversions among the Arabidopsis actin genes is somewhat surprising. However, DROUIN and DOVER (1990) also found no evidence for gene conversion among several potato actin genes. Taken together, our results imply that, in general, gene conversion is not a significant force in plant actin evolution. It is possible that, although plant actin genes do share high levels of similarity, they contain subtle, but important variations. This variability could occur at only one or a few positions affecting a critical protein-protein interaction, as noted below. If the different genes do contain advantageous differences (*i.e.*, they encode required, functionally distinct proteins), then one would expect that selection against gene conversion would be quite strong.

Evolution of the Arabidopsis actin gene family: Preliminary studies of actin gene evolution in soybean, rice, and potato suggested the existence of ancient plant actin subclasses (DROUIN and DOVER 1990; MCELROY *et al.* 1990; THANGAVELU *et al.* 1993). Our analysis of synonymous and nonsynonymous substitutions, as well as the branching order of the actin gene geneologies in Figures 4 and 5, also suggest that the Arabidopsis actin genes have arisen from a series of gene duplications that have been occurring throughout vascular plant evolution. In both neighbor joining and parsimony trees, five of the six Arabidopsis subclasses are more closely related to genes from other species than they are to their own family members. *ACT7*, for example, is part of a clade that contains several dicot sequences and also a conifer actin sequence. This suggests that the initial divergence in Arabidopsis actin gene evolution occurred before the divergence of gymno-

sperms (Upper Devonian, 370 mya) (CRONQUIST 1981). The *ACT2/8* subclass is more closely related to two monocot actin genes than it is to *ACT7*, suggesting that the *ACT2/8* and *ACT7* lineages split before the monocot/dicot divergence. The proximity of the *ACT11* lineage to a rice actin gene suggests that it also diverged around the time of the monocot/dicot split. The *ACT1/3* and *4/12* subclasses appear to be the most recently diverged Arabidopsis actin subclasses and may have split after the divergence of the Dilleniidae, ~70–100 mya.

Additional support for an ancient origin of the Arabidopsis actin subclasses is provided by the amount of nonsynonymous substitutions in their coding sequences. HIGHTOWER and MEAGHER (1985) estimated that plant actins evolve at a similar overall rate to other eukaryotic actins (1% RNS/100 million years, ±50 million years). Because it is quite possible that plant actins are evolving more quickly than other actins, and based on the minimum value in this estimate, we calculated a divergence time of 150 mya for the divergence of the most closely related subclasses (*ACT1/3 vs. ACT4/12*, 3% RNS) and a divergence of 350 mya for the most distantly related subclasses (*ACT1/3* and *4/12 vs. ACT2/8*). The divergence times that we calculated for the Arabidopsis actins are consistent with the divergence times estimated from the branching order of the gene trees that included other plant actin sequences. Undoubtedly, the divergence of most of the actin subclasses (with the exception of the event separating *ACT1/3* and *ACT4/12*) predate major events in land plant evolution, such as the angiosperm radiation and the monocot-dicot split.

The most recent gene duplications were the ones that produced the closely related pairs: *ACT1* and *3*, *ACT2* and *8*, and *ACT4* and *12*. Each of these gene pairs differed by only one derived amino acid substitution in their encoded proteins and by 60.6, 55.7, and 83.5% silent nucleotide substitutions (SNS) at third codon positions, respectively. Based on a rate of 1% SNS/2 million years, which seems to be reasonably accurate for a number of plant and animal genes (MEAGHER *et al.* 1989; WOLFE *et al.* 1989), we estimate that these gene pairs diverged 25–40 mya. A number of mechanisms exist for gene duplication, the most common of which appears to be unequal crossing-over. It is likely that the *ACT5/9* pair arose in this way, because they are very closely linked. The other gene pairs may have arisen by a similar mechanism; however, a second intriguing possibility is that they were generated by a more "global" genome duplication during the evolution of the Brassicaceae 25–40 mya. This idea is supported by the roughly equivalent amount of SNS between the three actin pairs and between other Arabidopsis gene pairs that have been isolated in this lab (R. MEAGHER and J. SENECOFF, unpublished results). In addition, a recent paper (KOWALSKI *et al.* 1994) presents evidence for partial duplication of two Arabidopsis chromosome segments. We are currently investigating this hypothesis by determining the relative map positions of the members in each pair.

We have discussed two different criteria (nucleotide substitution rates and phylogenetic analyses) that support an ancient origin for the Arabidopsis actin subclasses. The regulatory patterns of the Arabidopsis actin genes also suggest that the subclasses are evolutionarily divergent, as will be discussed below. We had hoped to determine in this study whether the Arabidopsis actin subclasses represented actin gene subclasses that are conserved in all higher plants. The tree topologies suggest that this is indeed the case; however, bootstrap support for many of the clades in the trees was unexpectedly low. Thus, although the weight of the evidence supports an ancient origin for plant actin gene subclasses, we cannot confidently designate orthologues of the Arabidopsis genes in other species with the data at hand. We expect that this will change as more actin gene sequences from other plants become available. The current data set contains only limited samples from gene families that are much larger than that of Arabidopsis. It is likely that conserved plant actin subclasses in these other species are incompletely represented due to sampling error, and most may not be represented at all. As more plant actin gene families are characterized, particularly from nonangiosperms, the resolution of plant actin gene phylogenies should increase. An intriguing possibility is that other plants contain unique actin gene lineages that fulfill important functional roles but that are not present in Arabidopsis. This study, in combination with functional studies of the Arabidopsis genes, should lead to hypotheses linking plant actin gene evolution and function that can be tested by a comparative approach in other plants.

Divergence of protein structures among Arabidopsis actin subclasses: Amino acid interchanges that would alter charged residues or peptide backbone structure were observed among the eight functional Arabidopsis actins. This variation appears far more significant than that found among the six human actins (Figure 7). The eight variable charged positions among the six human actins showed only conservative interchanges (*e.g.*, asn/gln or asp/glu). In contrast, only two of the nine positions that contained variation in charged residues among the eight highly expressed Arabidopsis actins were conservative. The other seven positions showed a variety of nonconservative substitutions (see RESULTS). In addition, two interchanges between standard α-amino acids and the imino acid proline were observed. Most of these amino acid changes supported the gene tree presented for Arabidopsis actin in that the most closely related genes ususally contained the same residues while the differences were observed among the most distant members of the family. For example, the *ACT2* and *ACT8* proteins contain a charged and a polar residue at positions Asp5 and Asn50, respectively, while the remaining six functional Arabidopsis actins did not.

Ser79 in the *ACT2* and *ACT8* subclass was shared by the next two most closely related subclasses *ACT7* and *ACT11,* but was not found in the more distant subclasses represented by *ACT1,* ACT3, *ACT4,* and *ACT12.* The importance of charged residue interchanges in actin function is supported by site-directed mutagenesis in which alanines were substituted for adjacent charged residues in the single yeast actin protein (WERTMAN *et al.* 1992). Most di-alanine substitutions had a strong phenotype or were lethal. Similarly, interchanges between amino and immune acids and their concomitant affects on the peptide backbone should be significant in a protein as conserved as actin. The presence of pro7 in all plant actins appears to alter its properties relative to animal and yeast actins, which have an alanine in this position. The substutition of pro7 in plant actins caused a structural rearrangement, relative to the animal protein. Preliminary data suggest that when this proline residue is substituted into the animal actin backbone and expressed in yeast, the resulting protein has a 10-fold lower critical concentration for polymerization (AL-ALMI *et al.* 1993).

It is perhaps significant that the nonconservative substitutions among the plant proteins were located on several surfaces of the actin molecule. They may have the potential to alter actin-actin or actin-actin binding protein interactions. For some of the variable residues there is direct evidence that this should be the case. For example, the his41 in *ACT2* and *ACT8* replaces the thr41 in the other functional plant actins and a gln41 in most animal actins. Gln41 projects right into one of the two main areas involved in actin monomer polymerization and myosin binding (KASPRZAK *et al.* 1988; SHETERLINE and SPARROW 1994). The ser170 present in *ACT2* and *ACT8* replaces the ala170 found in most plant and animal actins. This appears to be a primary contact with α-actinin (SHETERLINE and SPARROW 1994). The binding of profilin and gelsolin to actin may also be affected by a substitution of ser170 in plant actin since the adjacent residues phe/tyr169 and leu171 are contact points for these proteins. Even the conservative replacements of charged residues in the amino terminal peptide (asp1/glu1 and asp3/glu3) among the Arabidopsis actins might affect polymerization and myosin binding, since this domain is involved in polymerization (BERTRAND *et al.* 1989; SHETERLINE and SPARROW 1994).

Evolution of Arabidopsis actin gene regulation: As discussed above, the divergence in the *Arabidopsis* proteins supports the hypothesis that they have specialized functions. The second part of our two-part working hypothesis suggests that plant actin gene subclasses were preserved because they have evolved differential expression patterns. Companion studies from our laboratory, involving actin gene-specific mRNA analyses and promoter-reporter fusion expression in transgenic plants, demonstrate that this is indeed the case, as summarized in Figure 8. Five of the six subclasses are regulated in distinctly different ways and one is not expressed at detectable levels. Interestingly, the tissue- and organ-specific expression patterns of the five active actin subclasses are congruent with the evolutionary relationships based on coding sequence (*i.e.,* the most closely related genes have the most similar regulatory patterns, while the most distant subclasses have the least similar patterns). For example, the members of each closely related gene pair have almost identical expression patterns. *ACT1* and *3* are both expressed in mature pollen and all organ primordia, while *ACT4* and *12* are expressed in mature pollen and young vascular tissue. These two subclasses, which are the most closely related of the six, also have the most similar regulation. *ACT11* is expressed preferentially in pollen, ovules, and developing embryos. These five genes are most strongly expressed in reproductive organs, and thus constitute what we have defined as the reproductive actin class, although many of them also are active in some vegetative structures (*e.g., ACT11* is strongly expressed in the leaves and stem of the inflorescence). At the other end of the phylogeny are three genes that we have placed in a vegetative class. *ACT2* and *8* have a dramatically different expression pattern from any of the above genes in the reproductive class: they are expressed strongly and constitutively in most vegetative tissues but weakly or not at all in pollen or ovules. *ACT7* is the most similar subclass to *ACT2/8* in the sense that it is expressed in most vegetative tissues; however, it is preferentially expressed in younger, rapidly developing

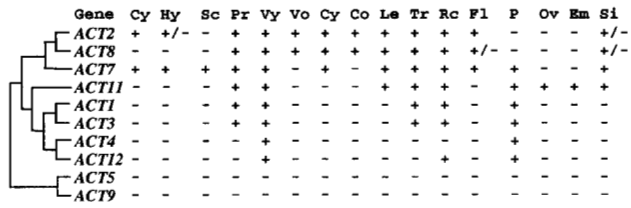

FIGURE 8.—Summary of the expression patterns of the Arabidopsis gene actin genes. The phylogenetic relationships of the 10 Arabidopsis actin genes, based on protein coding sequence, is presented on the left. To the right is a table that outlines the expression of the genes in the following tissues and organs: Cy, tissues of cotyledons excluding young vascular tissue; Hy, hypocotyl; Sc, seed coat; Pr, organ primordia and meristems; Vy, young vascular tissue; Vo, old vascular tissue; Cy, young cortical tissue; Co, old cortical tissue, Le, leaf epidermis; Rc, root cap; P, pollen; Ov, developing ovules; Em, embryo; Fl, flower (*i.e.,* sepals, petals, stigma, style) but excluding pollen, pollen sack, developing ovules; and Si, siliques. The + and − score expression and no expression, respectively, and a +/− score implies very weak expression or that expression is limited to a small subset of the cells present in this organ or tissue. These data are summarized from several recent manuscripts from this laboratory (Y.-Q.AN, S. HUANG, J. M. MCDOWELL, E. C. MCKINNEY and R. B. MEAGHER; Y.-Q. AN, J. M. MCDOWELL, S. HUANG, E. C. MCKINNEY, S. CHAMBLISS and R. B. MEAGHER; S. HUANG, Y.-Q. AN, J. MCDOWELL, E. C. MCKINNEY and R. B. MEAGHER; J. MCDOWELL, Y.-Q. AN, E. C. MCKINNEY, S. HUANG and R. B. MEAGHER; J. MCDOWELL, S. HUANG, E. C. MCKINNEY, Y.-Q. AN and R. B. MEAGHER) and is based on an analysis of RNA steady state levels and of the expression of promoter/reporter fusions in transgenic plants.

tissues and is the only Arabidopsis actin expressed in the hypocotyl and seed coat. While the phylogenetic split in the gene tree that separates the reproductive and vegetative classes of actin is fairly distinct, the details of their expression patterns reveal some overlap, suggesting the evolution of their regulatory elements may be even more complex that of their coding sequences (see below).

It should be noted that collectively these eight expressed genes account for strong actin expression in every tissue and organ of seedlings, juvenile and mature plants, and reproductive structures. This is consistent with our data above suggesting that all the actin gene family members have been isolated from Arabidopsis. It is of interest that the pollen-specific TAc25 gene is a close phylogenetic relative of the Arabidopsis actin genes that are preferentially expressed in pollen, supporting the notion that this is an ancient clade of actin genes. Further detailed characterizations actin gene expression in other plants may reveal other conserved regulatory subclasses, as has been seen in animal actin gene families.

There are at least two very different interpretations of how the complex patterns of actin gene regulation evolved, and both assume that duplicate actin gene copies are readily available. Based on a combinatorial-cassette view of promoter structure, tissue specific expression occurs through the interaction of multiple cis regulatory elements and the transacting factors that bind to them (DICKINSON 1988). One possible scenario is that the ancestral plant actin gene had all the possible regulatory elements for expression in all tissues. *ACT2* and *ACT8* have retained the elements for expression in vegetative tissues and lost most of those for expression in hypocotyl, seed cost, anther, and ovules. The other genes have retained elements for expression in pollen and other reproductive tissues, while losing selected vegetative elements. It is also likely that some ancient regulatory elements aquire new meanings as novel tissues and organs evolve. A second scenario is that each gene could accumulate novel regulatory elements as it evolves (BUREAU and WESSLER 1994; MCDONALD 1990). Considering the relatively recent origin of some tissues and organs in vascular plants (BREMER *et al.* 1987; CRANE 1990) and the very ancient character of at least portions of the plant actin gene family, the latter proposal seems more likely. Although the available data do not discriminate clearly between these possibilities, the experimental results we have obtained in Arabidopsis can be used to form testable hypotheses concerning the evolution of plant actin gene regulation.

The extreme divergence in expression patterns among the six subclasses also supports their hypothesized ancient origin, and suggests that differential expression is a significant raison d'être for plant actin multiplicity. However, considering the complexity of the evolutionary process, it would be premature to assign any causal relationships to either the differential regulation or the divergence in protein structure and function and the structure of the present-day actin gene family phylogeny. The combined role that selection, neutral drift, and contingency might play in linking actin gene family evolution and macroevolution has been considered separately (MEAGHER 1994). We hope that future studies examining Arabidopsis actin mutants (MCKINNEY *et al.* 1995) can be used to distinguish the importance of gene function and regulation to the maintenance of a complex plant actin gene family.

We acknowledge the comments of KATHERINE SPINDLER, MARCUS FECHHEIMER, and ROBERT PRICE on this manuscript. We also thank Dr. ROBERT FERL for kindly supplying his *ACT1* clone. This work was supported by a grant from the National Institute of Health Molecular Cytology Study Section. J.M.M. was supported by a National Research Service Award Predoctoral Training Grant fellowship in Genetics and this manuscript serves a partial fulfillment of the requirements for his Ph.D. in genetics.

LITERATURE CITED

AL-ALMI, J., C. MILLER, E. REISLER and P. A. RUBENSTEIN, 1993 The functional significance of actin Ala_7 (Abstr.). Am. Soc. Cell Biol. 33rd Annu. Meeting. pp. 208.

AN, Y.-Q., S. HUANG, J. M. MCDOWELL, E. C. MCKINNEY and R. B. MEAGHER, 1996 Conserved expression of the *Arabidopsis ACT1* and *ACT3* actin subclass in organ primordia and mature pollen. Plant Cell (in press).

BAIRD, W. V., and R. B. MEAGHER, 1987 A complex gene superfamily encodes actin in petunia. EMBO J. **6:** 3223–3231.

BERTRAND, R., P. CHAUSSEPIED, E. AUDERMARD and R. KASSAB, 1989 Functional characterization of skeletal F-actin labeled on the NH2-terminal segment of residues 1–28. Eur. J. Biochem. **181:** 747–754.

BREMER, K., C. J. HUMPHRIES, B. D. MISHLER and S. P. CHURCHILL, 1987 On cladistic relationships in green plants. Taxon **36:** 339–349.

BUREAU, T. E., and S. R. WESSLER, 1994 Mobile inverted-repeat elements of the tourist family are associated with the genes of many cereal grasses. Proc. Natl. Acad. Sci. USA **91:** 1411–1415.

CALLIS, J., T. CARPENTER, C.-W. SUN and R. D. VIERSTRA, 1995 Structure and evolution of the genes encoding polyubiquitin and ubiquitin-like proteins in *Arabidopsis thaliana* ecotype Columbia. Genetics **139:** 921–939.

CLARK, S. W., and D. I. MEYER, 1992 Centractin is an actin homologue associated with the centrosome. Nature **359:** 246–250.

CRANE, P. R., 1990 The phylogenetic context of microsporogenesis, pp. 11–14 in *Microspores: Evolution and Ontogeny*, edited by S. BLACKMORE and R. B. KNOX. Academic Press, New York.

CRONQUIST, A., 1981 *An Integrated System of Classification of Flowering Plants.* Columbia University Press, New York.

DERR, L. K., and J. N. STRATHERN, 1991 RNA-mediated recombination in *S. cerevisiae.* Cell **67:** 355–364.

DICKINSON, W. J., 1988 On the architecture of regulatory systems: evolutionary insights and implications. BioEssays **8:** 204–208.

DROUIN, G., and G. A. DOVER, 1987 A plant processed pseudogene. Nature **328:** 557–558.

DROUIN, G., and G. A. DOVER, 1990 Independent gene evolution in the potato actin gene family demonstrated by phylogenetic procedures for resolving gene conversions and the phylogeny of angiosperm actin genes. J. Mol. Evol. **31:** 132–150.

ELEFTHERIOU, E. P., and B. A. PALEVITZ, 1992 The effect of cytochalasin D on preprophase band organization in root tip cells of *Allium.* J. Cell Sci. **103:** 989–998.

FEINBERG, A. P., and B. VOGELSTEIN, 1983 A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:** 6–13 and Addendum.

FELSENSTEIN, J., 1993 PHYLIP (Phylogeny Inference Package), Univ. of Washington, Seattle, WA.

HALDANE, J. B. S., 1932 *The Causes of Evolution.* Longmans Green, London.

HERMAN, I. M., 1993 Actin isoforms. Curr. Opin. Cell Biol. **5:** 48–55.

HIGHTOWER, R. C., and R. B. MEAGHER, 1985 Divergence and differential expression of soybean actin genes. EMBO J. **4:** 1–8.

HIGHTOWER, R. C., and R. B. MEAGHER, 1986 The molecular evolution of actin. Genetics **114:** 315–332.

HIGUCHI, R. G., and H. OCHMAN, 1989 Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. Nucleic Acids Res. **17:** 5865.

HUANG, S., Y.-Q. AN, J, M. MCDOWELL E. C. MCKINNEY and R. B. MEAGHER, 1996 The *Arabidopsis thaliana ACT4/ACT12* actin gene subclass is strongly expressed in post-mitotic pollen. Plant J. (in press).

ITO, Y., S. ABE and E. DAVIES, 1994 Colocalization of cytoskeleton proteins and polysomes with a membrane fraction from pea. J. Exp. Bot. **45:** 253–259.

KABSCH, W., H. G. MANNHERZ, D. SUCK, E. F. PAI and K. C. HOLMES, 1990 Atomic structure of the actin: DNase I complex. Nature **347:** 37–44.

KASPRZAK, A. A., R. TAKASHI and M. F. MORALES, 1988 Orientation of actin monomer in the F-actin filament: radial coordinate of glutamine-41 and effect of myosin subfragment 1 binding on the monomer orientation. Biochemistry **27:** 4512–4522.

KOBAYASHI, H., H. FUKUDA and H. SHIBAOKA, 1987 Reorganization of actin filaments associated with the differentiation of tracheary elements in *Zinnia* mesophyll cells. Protoplasma **138:** 69–71.

KOWALSKI, S. P., T.-H. LAN, K. A. FELDMANN and A. H. PETERSON, 1994 Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. Genetics **138:** 499–510.

LEES-MILLER, J. P., D. M. HELFMAN and T. A. SHROER, 1992a A vertebrate actin-related protein is a component of a multisubunit complex involved in microtubule-based actin motility. Nature **359:** 244–246.

LEES-MILLER, J. P., G. HENRY and D. M. HELFMAN, 1992b Identification of *act2*, an essential gene in the fission yeast *Schizosaccharomyces pombe* that encodes a protein related to actin. Proc. Natl. Acad. Sci. USA **89:** 80–83.

LI, W.-H., C.-I. WU and C.-C. LUO, 1986 A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon usage. Mol. Biol. Evol. **2:** 150–174.

LLOYD, C. W., and J. A. TRAAS, 1988 The role of F-actin in determining the division plane of carrot suspension cells. Development **102:** 211–221.

MANIATIS, T., E. F. FRITSCH and J. SAMBROOK, 1989 *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

MASCARENHAS, J. P., 1993 Molecular mechanisms of pollen tube growth and differentiation. Plant Cell **5:** 1303–1314.

MCDONALD, J. F., 1990 Macroevolution and retroviral elements. Bio. Sci. **30:** 183–191.

MCELROY, D., M. ROTHENBERG, K. S. REECE and R. WU, 1990 Characterization of the rice *(Oryza sativa)* actin gene family. Plant Mol. Biol. **15:** 257–268.

MCKINNEY, E. C., N. ALI, A. TRAUT, K. A. FELDMANN, D. A. BELOSTOTSKY, J. A. MCDOWELL and R. B. MEAGHER, 1995 Sequence based identification of T-DNA insertion mutations in *Arabidopsis:* actin mutants *act2-1* and *act4-1.* Plant J. **8:** 613–622.

MCLEAN, M., W. V. BAIRD, A. G. M. GERATS and R. B. MEAGHER, 1988 Determination of copy number and linkage relationships among five actin gene subfamilies in *Petunia hybrida.* Plant Mol. Biol. **11:** 663–672.

MCLEAN, B. G., S. HUANG, E. C. MCKINNEY and R. B. MEAGHER, 1990 Plants contain highly divergent actin isovariants. Cell Motil. **17:** 276–290.

MEAGHER, R. B., 1991 Divergence and differential expression of actin gene families in higher plants. Int. Rev. Cytol. **125:** 139–163.

MEAGHER, R. B., 1994 The impact of historical contingency on gene phylogeny: plant actin diversity, pp. in *Evolutionary Biology,* Vol. 28, edited by M. K. HECHT, R. J. MACINTYRE, and M. T. CLEGG. Plenum Press, New York.

MEAGHER, R. B., S. BERRY-LOWE and K. RICE, 1989 Molecular evolution of the small subunit of ribulose bisphosphate carboxylase: nucleotide substitution and gene conversion. Genetics **123:** 845–863.

MEAGHER, R. B., and R. E. WILLIAMSON, 1994 The Plant cytoskeleton, pp. 1049–1084 in *Arabidopsis,* edited by E. MEYEROWITZ and C. SOMERVILLE. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

MINEYUKI, Y., and B. A. PALEVITZ, 1990 Relationship between preprophase band organization, F-actin and the division site in *Allium.* Fluorescence and morphometric studies on cytochalasin-treated cells. J. Cell Sci. **97:** 283–295.

MIWA, T., K. MANABE, K. KUROKAWA, S. KAMADA, N. KANDA *et al.,* 1991 Structure, chromosome location, and expression of the human smooth muscle (enteric type) gamma-actin gene: evolution of six human actin genes. Mol. Cell. Biol. **11:** 3296–3306.

NAGAO, R. T., D. M. SHAH, V. K. ECKENRODE and R. B. MEAGHER, 1981 Multigene family of actin-related sequences isolated from soybean genomic library. DNA **2:** 1–9.

OHNO, S., 1970 Why gene duplication? pp. 59–65 in *Evolution by Gene Duplication.* Springer-Verlag, New York.

RUSSELL, S. D., 1993 The egg cell: development and role in fertilization and early embryogenesis. Plant Cell **5:** 1349–1359.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. J. SCHARF, R. HIGUCHI *et al.,* 1988 Primer-direced enzymataic amplificaion of DNA with a thermostable DNA polymerase. Science **239:** 487–491.

SCHWOB, E., and R. P. MARTIN, 1992 New yeast actin-like gene required late in the cell cycle. Nature **355:** 179–182.

SEAGULL, R. W., 1990 The effects of microtubule and microfilament disrupting agents on cytoskeletal arrays and wall deposition in developing cotton fibers. Protoplasma **159:** 44–59.

SHAH, D. M., R. C. HIGHTOWER and R. B. MEAGHER, 1983 Genes encoding actin in higher plants: intron positions are highly conserved but the coding sequences are not. J. Mol. Appl. Gen. **2:** 111–126.

SHETERLINE, P., and J. C. SPARROW, 1994 Actin. Protein Profile **1:** 1–62.

SWOFFORD, D. L., 1990 PAUP: Phylogenetic analysis using parsimony, version 3.1, Univ. of Illinois, Champaign, IL.

THANGAVELU, M., D. BELOSTOTSKY, M. W. BEVAN, R. B. FLAVELL, H. J. ROGERS *et al.,* 1993 Partial characterization of the *Nicotiana tabacum* actin gene family: evidence for pollen specific expression of one of the gene family members. Mol. Gen. Genet. **240:** 290–295.

VOYTAS, D. F., and F. M. AUSUBEL, 1988 A copia-like transposable element family in *Arabidopsis thaliana.* Nature **336:** 242–244.

WERNICKE, W., and G. JUNG, 1991 Role of the cytoskeleton in cell shaping of developing mesophyll of wheat (Triticum aestivum L.). Eur. J. Cell Biol. **57:** 88–94.

WERTMAN, K. F., D. G. DRUBIN and D. BOTSTEIN, 1992 Systematic mutational analysis of the yeast *ACT1* gene. Genetics **132:** 337–350.

WILLIAMSON, R. E., 1993 Organelle movements. Annu. Rev. Plant Physiol. Plant Mol. Biol. **44:** 181–202.

WOLFE, K. H., P. M. SHARP and W.-H. LI, 1989 Rates of synonymous substitution in plant nuclear genes. J. Mol. Evol. **29:** 208–211.

Communicating editor: J. CHORY