THE MOLECULAR EVOLUTION OF ACTIN

ROBIN C. HIGHTOWER[1] AND RICHARD B. MEAGHER

Department of Genetics, University of Georgia, Athens, Georgia 30602

Manuscript received January 6, 1986
Revised copy accepted May 20, 1986

ABSTRACT

We have investigated the molecular evolution of plant and nonplant actin genes comparing nucleotide and amino acid sequences of 20 actin genes. Nucleotide changes resulting in amino acid substitutions (replacement substitutions) ranged from 3–7% for all pairwise comparisons of animal actin genes with the following exceptions. Comparisons between higher animal muscle actin gene sequences and comparisons between higher animal cytoplasmic actin gene sequences indicated <3% divergence. Comparisons between plant and nonplant actin genes revealed, with two exceptions, 11–15% replacement substitution. In the analysis of plant actins, replacement substitution between soybean actin genes *SAc1*, *SAc3*, *SAc4* and maize actin gene *MAc1* ranged from 8–10%, whereas these members within the soybean actin gene family ranged from 6–9% replacement substitution. The rate of sequence divergence of plant actin sequences appears to be similar to that observed for animal actins. Furthermore, these and other data suggest that the plant actin gene family is ancient and that the families of soybean and maize actin genes have diverged from a single common ancestral plant actin gene that originated long before the divergence of monocots and dicots. The soybean actin multigene family encodes at least three classes of actin. These classes each contain a pair of actin genes that have been designated kappa (*SAc1*, *SAc6*), lambda (*SAc2*, *SAc4*) and mu (*SAc3*, *SAc7*). The three classes of soybean actin are more divergent in nucleotide sequence from one another than higher animal cytoplasmic actin is divergent from muscle actin. The location and distribution of amino acid changes were compared between actin proteins from all sources. A comparison of the hydropathy of all actin sequences, except from Oxytricha, indicated a strong similarity in hydropathic character between all plant and nonplant actins despite the greater number of replacement substitutions in plant actins. These protein sequence comparisons are discussed with respect to the demonstrated and implicated roles of actin in plants and animals, as well as the tissue-specific expression of actin.

MULTIGENE families are sets of genes descended by duplication and divergence from common ancestral genes. These families are useful models for studying the evolution of eukaryotic genes in relation to their function and regulation. Comparisons of the protein coding regions, nucleotide sequences and exon-intron arrangements of related genes provide a means of

[1] Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143.

tracing their evolutionary pathways (EFSTRATIADIS *et al.* 1980; BROWN *et al.* 1984).

The structure of actin genes can be examined across broad evolutionary distances because actin is highly conserved and is ubiquitous in eukaryotes. In animals, actin is primarily involved in muscle contraction in differentiated muscle tissue. In nonmuscle animal cells, actin is involved in a variety of processes including cytoskeletal structure, cellular motility, cell-surface mobility, intracellular transport, cytoplasmic streaming, cytokinesis, endocytosis, exocytosis, clot retraction, microvillar movement and, possibly, chromosomal condensation and mitosis (SCHLIWA 1981; LLOYD 1983; PONTE *et al.* 1983; STOSSEL 1984).

Actin is encoded by a multigene family in all animals, and protozoa and plants so far examined, but is encoded by a single gene in yeast. The organization of this gene family from animals and protozoa has been studied in Dictyostelium (MCKEOWN and FIRTEL 1981), *Caenorhabditis elegans* (FILES, CARR and HIRSH 1983), Drosophila (FYRBERG *et al.* 1981), *Strongylocentrotus purpuratus* (DAVIDSON *et al.* 1982), chicken (FORNWALD *et al.* 1982), rat (ZAKUT *et al.* 1982) and human (ENGEL, GUNNING and KEDES 1981). The gene family differs in size among these organisms, and the number and location of introns within actin genes can be highly variable (SHAH, HIGHTOWER and MEAGHER 1983).

Slight sequence differences located primarily in the amino terminal region of mammalian actins result in six distinct actin isoforms (VANDEKERCKHOVE and WEBER 1978). These differences identify three major categories of actin in mammals: nonmuscle or cytoplasmic types (β and γ in both brain and thymus), smooth muscle types (vascular and nonvascular) and striated muscle types (α skeletal and α cardiac). Smooth-muscle and striated-muscle actins are more closely related to each other than they are to cytoplasmic actins.

A limited number of studies have identified actin in plants (WILLIAMSON 1980). JACKSON and DOYLE (1977) identified a putative actin from *Phaseolus vulgaris* root tips. Actin has been isolated from wheat germ (ILKER, BREIDENBACH and MURPHY 1979), tomato endocarp tissue (VAHEY and SCORDILIS 1980) and soybean seedlings (METCALF *et al.* 1980). Plant actin isotypes have not yet been identified. The sequences of the first nine amino acids in the amino terminus of the three soybean actins and one maize actin, however, are far more conserved among distant plant actins than the corresponding sequences that result in isotypic variation in animal actins (SHAH, HIGHTOWER and MEAGHER 1983).

Despite the knowledge that actin is a fundamental component of animal cells, the role of actin in plants has not been so well characterized. Actin filaments generate the force required for cytoplasmic streaming in Nitella (HIGASHI-FUJIME 1980) and Chara (WILLIAMSON 1975). Plant actins appear to function in other motile processes in plant cells such as the movement of cellular organelles, in particular chloroplast aggregation (BLATT, WESSELLS and BRIGGS 1980) and rotation (KLEIN, WAGNER and BLATT 1980), as well as having potential roles in membrane movement and chromosome movement during mitosis and meiosis (LLOYD 1983).

Actin gene families in plants have been studied in soybean and maize (SHAH, HIGHTOWER and MEAGHER 1982, 1983; HIGHTOWER and MEAGHER 1985). The size of the gene families differs between these two species, but the number and location of introns within protein coding regions are conserved in those genes so far examined.

The purpose of the following study was to determine the evolutionary relationship of soybean actin to actins in other eukaryotes. The percentage of nucleotide substitution that results in amino acid replacement (replacement substitutions) between actin genes from plant and nonplant organisms is calculated. Nucleotide replacement substitution is the most accurate value for estimating the number of mutational events leading to amino acid change (PERLER *et al.* 1980). The distribution of replacement substitutions in actin gene sequences is examined.

Comparisons of hydropathy values of each actin protein are also presented in order to examine the effects of the resulting amino acid replacements on hydrophilic and hydrophobic domains of actin. Collectively, these data compare diverse actin genes and proteins from organisms within and among the protozoan, fungal, plant and animal kingdoms. These results are discussed in the context of the presumed roles of actin in plants and animals.

MATERIALS AND METHODS

Actin gene sequence data were stored and analyzed on a Bion Workstation (Bion Intelligenetics, Palo Alto, California). The majority of DNA sequence data was obtained through the National Institutes of Health database (NG and ABELSON 1980; COOPER and CRAIN 1982; FORNWALD *et al.* 1982; HAMADA, PETRINO and KAKUNAGA 1982; KAINE and SPEAR 1982; NELLEN and GALLWITZ 1982; SHAH, HIGHTOWER and MEAGHER 1982, 1983; ZAKUT *et al.* 1982; NUDEL *et al.* 1983; SANCHEZ *et al.* 1983; SCHULER, MCOSKAR and KELLER 1983), apart from recent publications, for which DNA sequence data were entered manually (PONTE *et al.* 1984; UEYAMA *et al.* 1984; BERGSMA, CHANG and SCHWARTZ 1985). The human γ actin sequence was kindly provided before publication by HARRY ERBA and LARRY KEDES of Stanford University. MIKE MCLEAN and TOM MCKNIGHT generously supplied the *SAc4* soybean actin gene sequence before publication. Twenty sequences from organisms representing four eukaryotic kingdoms were included in the analysis: protozoa: Acanthamoeba, Oxytricha; fungal: yeast; plant: soybean, maize; animal: chick, Drosophila, human, rat, sea urchin.

The percentage of silent and replacement substitutions between pairwise comparisons of actin genes was calculated as described by PERLER *et al.* (1980). Initial calculations were done on an Apple II plus computer using a program provided by EFSTRADIATIS and FULLER (SHAH, HIGHTOWER and MEAGHER 1983). All analyses presented here were done on the Bion workstation using an adaptation of this program, called DIV, which was written by KEN RICE (unpublished results). Graphics showing the distribution of replacement substitutions between pairwise gene comparisons were generated with a companion program developed by KEN RICE. A phenogram illustrating the evolutionary relationships of actin genes from plant, fungal and animal sources was generated with the percentage replacement substitution data (see Figure 2) by the program GMDP1M-cluster analysis of variables (BMDP Statistical Software, Inc., Los Angeles), based on the clustering strategy UPGMA (unweighted pair-group method using arithmetic averages) (SOKAL and SNEATH 1963). Hydropathy profiles, based on calculations of KYTE and DOOLITTLE (1982), were generated using the PEP software of Intelligenetics Co. To quantify the degree of similarity among hydropathy profiles, Pearson's product

moment correlation coefficient was calculated for each pairwise actin protein comparison (SAS INSTITUTE 1982).

A χ^2 test was used to determine if amino acid substitutions were randomly distributed throughout actin protein sequences. The sequences were divided into ten intervals, and observed and expected numbers of changes were recorded within each interval. At each position within an interval, one amino acid substitution was recorded whenever one or more changes were present in the 20 actin sequences.

RESULTS

The four plant actin genes examined, *SAc1*, *SAc3*, *SAc4* and *MAc1*, have diverged from nonplant actins in 25–30% of their amino-acid-encoding nucleotide sequences (except for Oxytricha) and approximately 22% from each other. Of the 377 amino acids compared, there were from 50 to 70 amino acid replacements between these plant actins and most of the nonplant actins, and from 30 to 50 amino acid replacements between each pairwise comparison of the four plant actins. To define and quantify the divergence of these plant actin genes and proteins from each other and from nonplant actin sequences, we have compared these sequences by a number of criteria. Because of the difficulty of aligning amino terminal sequence data, the following analyses began on data corresponding to amino acid residue 9 (alanine in all nonplant actins, proline in all plant actins) to optimize homology. Results, therefore, might be slightly underestimated or overestimated for each particular analysis. A figure of the amino acid alignment for the 20 sequences analyzed is available from the authors (R.B.M.) upon request.

Hydropathy profiles of actin sequences: The hydropathy profiles comparing the hydrophilicity and hydrophobicity of each actin protein sequence indicated very strong similarities among all sequences examined, with the exception of Oxytricha (to be discussed separately). Examples illustrating this similarity for a portion of the actin sequence (amino acids 20–152) are shown in Figure 1. Correlation coefficients ranged from 0.91 to 1.00 for all pairwise comparisons, with the exception of comparisons with Oxytricha actin, which ranged from 0.77 to 0.82. When these actin sequences were compared to a nonrelated sequence of similar length, the Tu factor (AN and FRIESEN 1980), the correlation coefficients were between 0 and 0.01.

Positive correlations would be expected for highly conserved proteins like actin. In addition to the high amino acid homology, most amino acid changes in actin proteins result in conservative amino acid replacements. The following comparisons were performed to evaluate the contribution of conservative amino acid replacements to the high positive correlation between actin proteins. The actin protein sequence encoded by *SAc3* was compared to the actin protein encoded by Acanthamoeba, chick α-skeletal, sea urchin 2, and *SAc1*. At all differing positions, the amino acid residues in *SAc3* were replaced by the least conservative amino acid residue relative to the corresponding residue in the protein being compared (KYTE and DOOLITTLE 1982). Correlation coefficients for the four pairwise comparisons with the modified *SAc3* hydropathy profile were 0.72, 0.68, 0.72 and 0.71, respectively, representing 43, 51, 45 and 47 amino acid changes. The correlations determined for unmodified *SAc3*

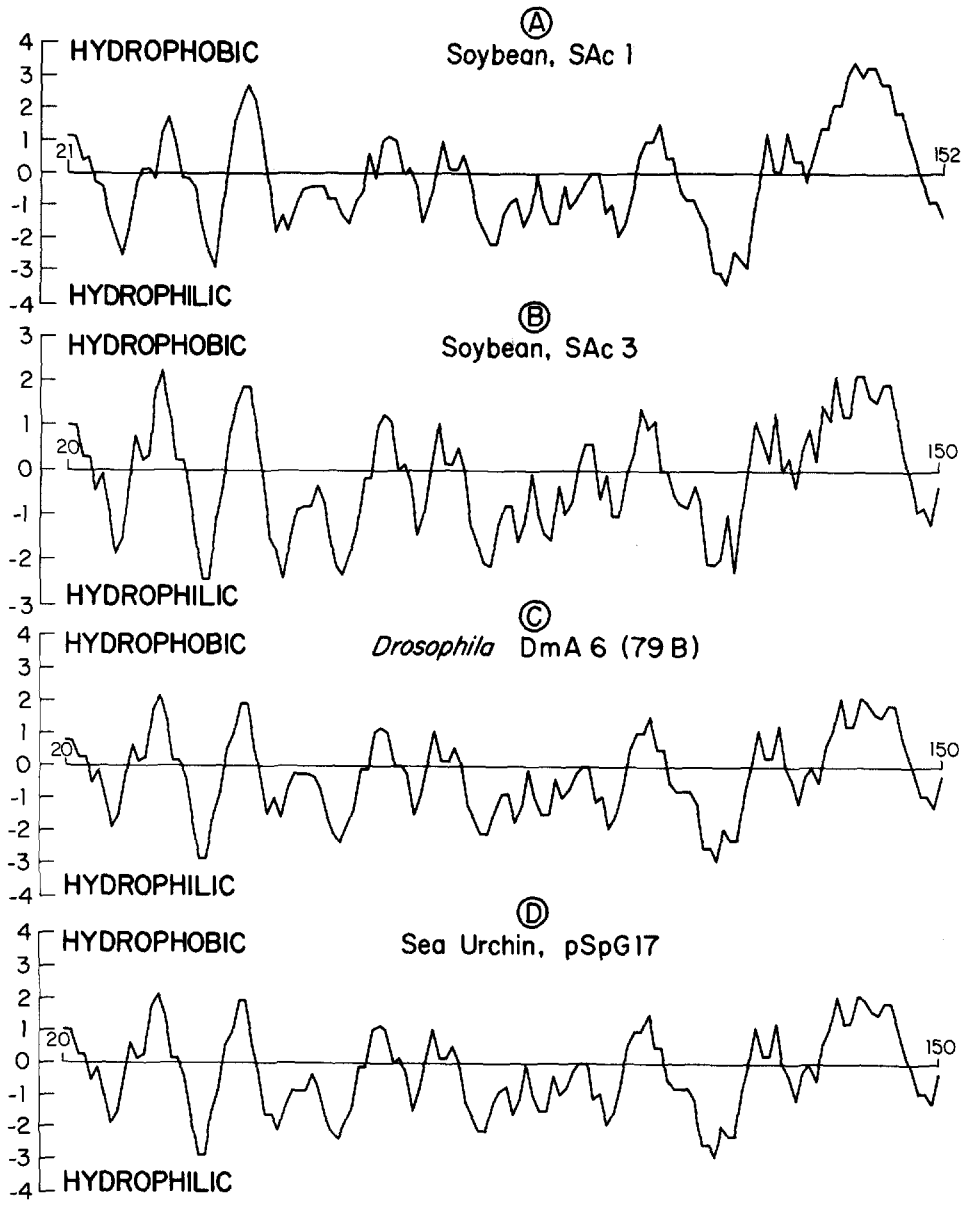

FIGURE 1.—Hydropathy profiles of actins from soybean, Drosophila and sea urchin spanning the region between amino acid residues 21–152. Hydrophobic areas are denoted by peaks above the base line (at 0), and hydrophilic regions are denoted by peaks below the base line. Note the very similar profiles among the four actins. Sea urchin pSpG17 refers to sea urchin 3, SU·3.

and the four sequences were 0.96, 0.93, 0.95 and 0.94, respectively. The relative differences in these values suggest that conservative amino acid replacements in actin proteins have contributed to the high positive correlations observed among them.

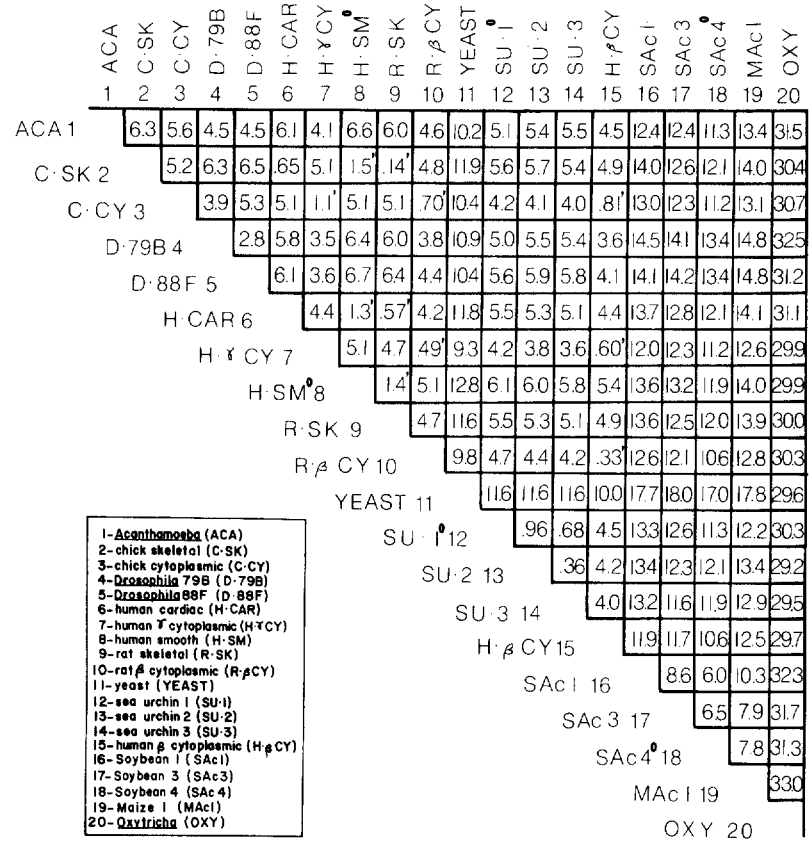

FIGURE 2.—Percentage of nucleotide replacement substitution between the actin coding sequences. Pairwise comparisons of animal muscle and nonmuscle actin genes with nucleotide replacement substitution values ranging from 0.14 to 1.5% are denoted by the symbol '. Muscle actin genes are identified as skeletal, cardiac or smooth. Nonmuscle actin genes are identified as cytoplasmic, β-cytoplasmic or γ-cytoplasmic. Incomplete sequences are indicated by the symbol °.

Silent nucleotide substitution: A large percentage of silent substitution was observed between plant-nonplant actin gene comparisons. After a correction for multiple-hit kinetics (PERLER *et al.* 1980), each plant-nonplant actin gene comparison revealed >100% silent substitution, suggesting that most silent sites have been substituted several times. This was also observed for plant-plant actin gene comparisons, which indicates that they also have been saturated for silent site substitutions since their divergence from a common ancestor. Therefore, we have focused on percentage replacement substitution (Figure 2), a more reliable measure of the gene divergence over long periods of time (LEWIN 1983).

Nucleotide replacement substitution: The nucleotide replacement substitutions between the three soybean actin genes *SAc1*, *SAc3*, and *SAc4* and the maize actin gene *MAc1* were 10%, 8% and 8%, respectively (Figure 2). Pairwise comparisons between *SAc1-SAc3*, *SAc1-SAc4* and *SAc3-SAc4* actin genes revealed nucleotide replacement substitution values of 9%, 6% and 6%, respectively.

The nucleotide replacement substitution between each of the four plant actin genes and yeast was 17–18%, between Oxytricha and each of the plant actin genes was 31–33% and between the four plant actin genes and all other nonplant actin genes considered individually was 11–15% (Figure 2).

The degree of replacement substitution ranged from 3 to 7% for all nonplant actin gene comparisons, with the following exceptions. Replacement substitution of <1% was observed for the three sea urchin actin genes examined. Percentage replacement substitutions for 11 pairwise comparisons of higher animal muscle actin genes with other muscle actin genes and of nonmuscle actin genes with other nonmuscle actin genes ranged from 0.14 to 1.5% (Figure 2). The highest replacement substitution, ranging from 29 to 33% (Figure 2), was found between Oxytricha and every other actin.

A phenogram, using percentage replacement substitution data (Figure 2), illustrates the possible evolutionary relationships of actin genes from plant, fungal and animal sources (Figure 3).

Distribution of amino acid substitutions: The location and relative frequency of nucleotide replacement substitution between various pairwise gene comparisons was plotted to illustrate conserved and variable regions in the actin gene sequences. Figure 4 presents four divergence plots comparing two soybean actin genes, a soybean and animal actin gene, a protozoan and an animal actin gene, and two animal actin genes. These three examples demonstrate that, among the most divergent eukaryotes examined, replacement substitutions were located essentially at random.

Results from a χ^2 test applied to the amino acid sequence comparisons of these sequences divided into ten intervals (not shown; MATERIALS AND METHODS) also indicated that amino acid substitutions were essentially randomly distributed throughout actin proteins when actins from all four kingdoms were compared ($\chi^2 = 11.1$; d.f. = 9, $P > 0.05$). However, the positions and distributions of three categories of amino acid substitutions were determined: those positions that varied in amino acid substitutions in (1) exclusively plant gene sequences, (2) exclusively nonplant gene sequences and (3) in both plant and nonplant gene sequences (Table 1). The amino acid residues in 55 positions were different in one to four of the plant actin genes, while being conserved in all other nonplant actin sequences examined. The amino acid residues in 43 positions were different in one to ten of the nonplant actin gene sequences examined, while being conserved in all plant actins. The amino acid residues in 34 positions were not conserved in either plant or nonplant actin gene sequences examined. Oxytricha was not included in this comparison.

DISCUSSION

Rates of nucleotide replacement substitution: Nucleotide substitutions that result in amino acid replacements accumulate at an overall linear rate during the evolution of many conserved genes (WILSON, CARLSON and WHITE 1977; PERLER *et al.* 1980). For example, globin genes in animals diverge at nucleotide replacement substitution rates of 1% per 10 million years (MY) (EFSTRATIADIS *et al.* 1980). If we compare replacement substitutions among the representative

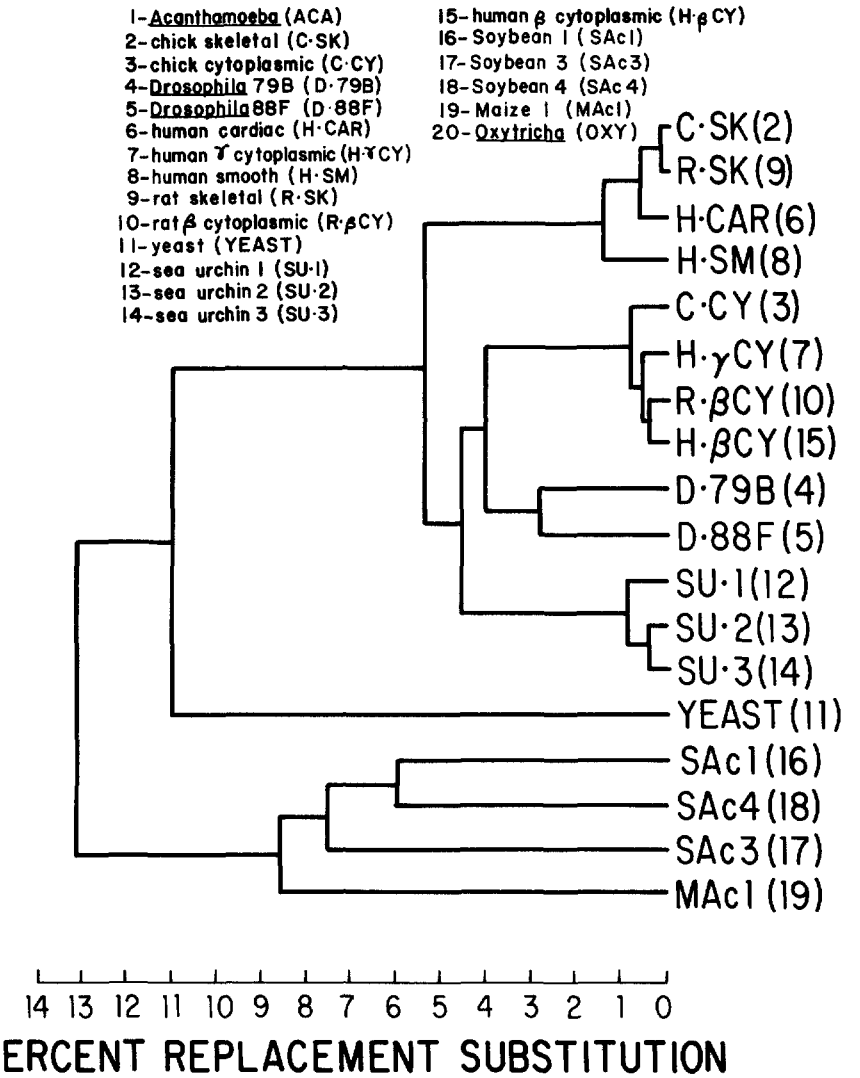

FIGURE 3.—Phenogram of actin genes from three eukaryotic kingdoms: the relationships of actin genes from organisms representing plant, fungal and animal kingdoms are displayed. A pairwise cluster analysis was used with the percentage replacement substitution data set (Figure 2) to evaluate the relationships based on the divergence among genes.

actin genes from the four eukaryotic kingdoms and assume a divergence time of 1000 MY for all kingdoms (MARGULIS and SCHWARTZ 1976), we estimate 1% per 30–244 MY for the overall rate of nucleotide replacement substitution. If the values for the protozoan actins (Acanthamoeba and Oxytricha) are omitted, then this range of rates for the remaining three kingdoms is much narrower, 1% per 56–108 MY (Table 2). The divergence rates for plant actin genes compared to actin genes from the animal and fungal kingdoms are 1% per 56–94 MY. Representatives of the protostome and deuterostome branches

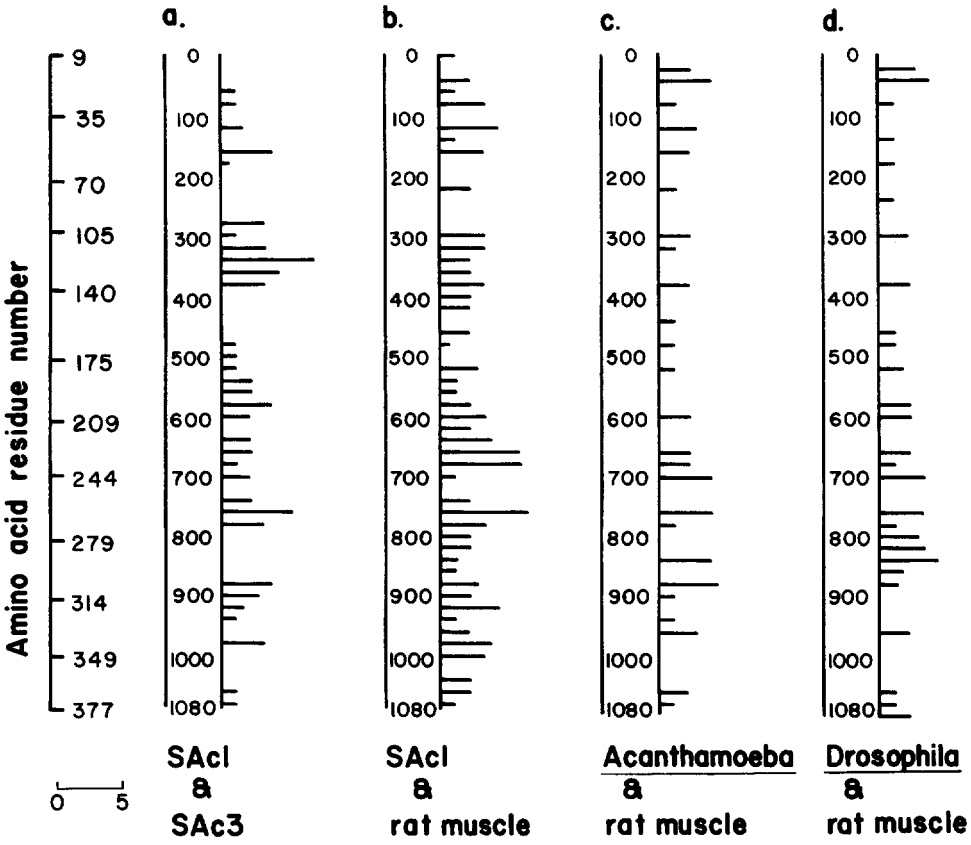

FIGURE 4.—Divergence plots. The relative frequency of nucleotide replacement substitutions between four pairwise comparisons of actin genes is plotted to illustrate the extent and distribution of nucleotide replacement substitutions. Plant-plant, plant-animal, protozoan-animal and animal-animal actin comparisons are shown. The amino acid residue numbers (1–377) corresponding to the nucleotide sequence numbers (1–1080) are indicated at the left margin of the figure. A bar representing five nucleotide replacement substitutions is shown at the lower left.

of the animal kingdom such as Drosophila and sea urchin, respectively, are thought to have diverged between 550 and 650 MY ago (MARGULIS and SCHWARTZ 1976). The nucleotide replacement substitution rates between actin genes from representatives of these branches of the animal kingdom are about 1% per 100 MY, within our calculated overall range of rates. We conclude from these data (Table 2) that the plant actin genes have diverged from non-plant actins at overall rates similar to those rates of divergence of most non-plant actin genes from each other (*i.e.*, animal *vs.* fungal or deuterosome *vs.* protostome). These actin data are plotted in Figure 5; the nucleotide replacement substitution rate for globin genes is included for comparison (EFSTRATIADIS *et al.*, 1980). It can be seen that most of the values for animal protostome *vs.* animal deuterostome actin comparisons (cluster 3) and plant *vs.* non-plant actin comparisons (cluster 2) fall within a strikingly narrow range, which represents between 0.9 and 1.8% per 100 MY. One would not expect a com-

TABLE 1

Three categories of amino acid changes occurring in actin proteins

I. Amino acid residues in the following positions are variable in plant actins only, conserved in nonplant actins (55 of 369 amino acids)
9, 10, 16, 23, 27, 46, 62, 77, 79, 86, 99, 100, 102, 103, 106, 107, 123, 124, 130, 135, 138, 139, 190, 197, 202, 205, 207, 209, 219, 221, 225, 229, 232, 233, 242, 244, 254, 258, 261, 264, 270, 273, 275, 276, 279, 309, 316, 320, 323, 331, 332, 335, 336, 339, 340

II. Amino acid residues in the following positions are variable in nonplant actins only, conserved in plant actins (43 of 369 amino acids)
12, 17, 18, 19, 45, 68, 70, 73, 87, 91, 112, 137, 146, 164, 169, 178, 180, 196, 203, 217, 227, 236, 259, 263, 265, 267, 271, 277, 278, 280, 281, 282, 289, 291, 294, 301, 312, 313, 319, 325, 327, 346, 352

III. Amino acid residues in the following positions are variable in both plant and nonplant actins (34 of 369 amino acids)
43, 54, 78, 105, 116, 131, 134, 155, 162, 171, 172, 201, 214, 230, 234, 237, 238, 262, 268, 269, 274, 297, 299, 305, 308, 310, 321, 326, 329, 360, 362, 367, 370, 374

TABLE 2

Unit evolutionary periods of replacement substitution in actin genes from plant and nonplant sources

Comparison	Replacement substitution (%)	UEP
Plant *vs.* animal:	10.6–14.8	68–94
Plant *vs.* fungal	17.0–18.0	56–59
Animal *vs.* fungal	9.3–12.8	78–108
Animal *vs.* protist:		
Acanthamoeba	4.1–6.6	152–244
Oxytricha	29.2–32.5	31–34
Fungal *vs.* protist:		
Acanthamoeba	10.2	98
Oxytricha	29.6	34
Plant *vs.* protist:		
Acanthamoeba	11.3–13.4	75–89
Oxytricha	31.3–33.0	30–32

Unit evolutionary period (UEP) refers to the time in million years (MY) required for the fixation of 1% changes between two-lines (PERLER *et al.* 1980).

pletely linear rate of nucleotide replacement substitution over distances of 500–1000 MY; however, these data suggest the evolution of these genes has occurred at generally linear rates. In comparison to globin genes, the rates of replacement substitution for actin genes are very low. These low rates suggest that strong selective constraints have been imposed on actin.

Although the majority of actin comparisons (89 of 128) fall within this range of replacement substitution rates, two sets of comparisons fall outside these limits. Based on these rates, all comparisons with the Oxytricha actin gene (cluster 1) showed twice the level of replacement substitution expected, and all comparisons of an Acanthamoeba actin gene with animal actin genes (cluster

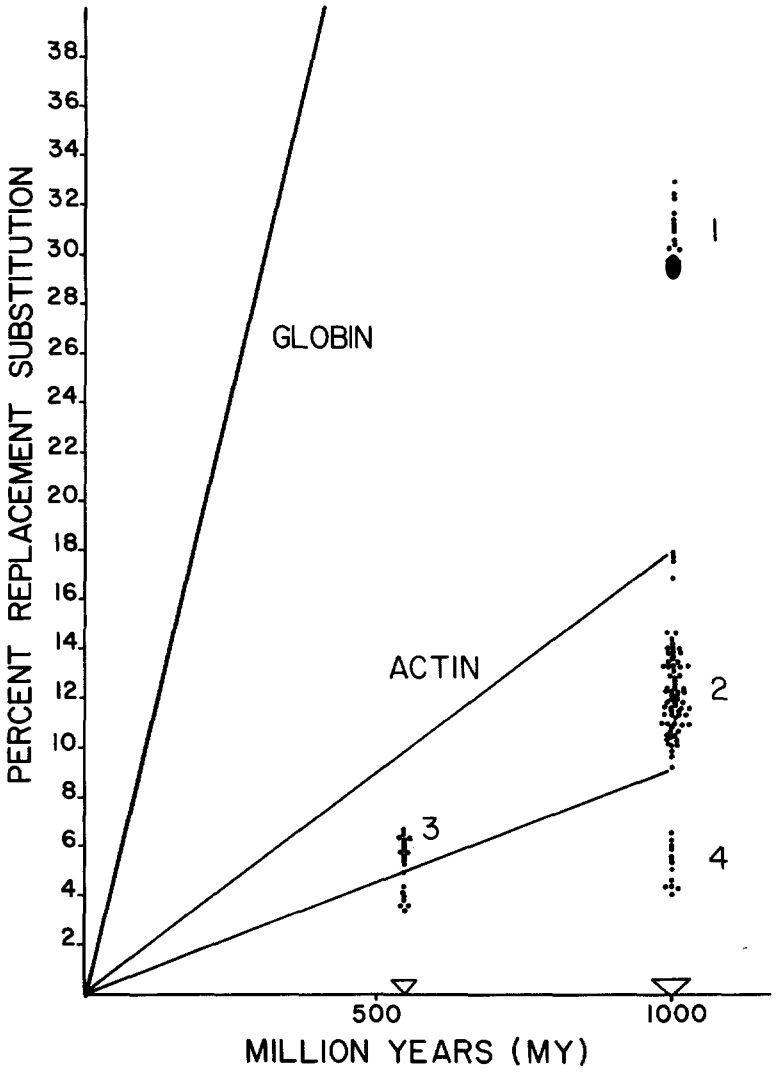

FIGURE 5.—Rate of nucleotide replacement substitution for representative actin genes from protozoan, fungal, plant and animal kingdoms. Percentage replacement substitutions for these comparisons were plotted at 1000 MY, the estimated divergence time for the eukaryotic kingdoms (MARGULIS and SCHWARTZ 1976). Comparisons of sequences from the representatives of the protostome branch (Drosophila) and the deuterostome branch (chicken, human, rat, sea urchin) of the animal kingdom have been plotted at 550 MY, the minimum estimated divergence time for these two branches (MARGULIS and SCHWARTZ 1976). Cluster one represents all comparisons with Oxytricha actin (19 points), cluster two contains all plant-nonplant and all fungal-animal actin comparisons (74 points), cluster three represents comparisons of actin genes from protostome and deuterostome lineages (22 points) and cluster four contains comparisons of an Acanthamoeba actin with animal actin genes (13 points). Plant-fungal and plant-animal actin comparisons are bracketed by two lines. The rate of replacement nucleotide substitution for globin genes is included for comparison (EFSTRATIADIS *et al.* 1980).

4) showed half the level expected (Figure 5). It is possible that the Oxytricha and Acanthamoeba actin genes have followed distinctly different rates of evolution from the other actin genes examined. These data could also be explained if the divergence times assumed for these organisms were extremely different from the currently assumed ones.

These comparisons of actin genes from the animal, fungal and plant kingdoms may be used to interpret the evolutionary relationships of the various soybean actin genes to one another. The three soybean actin genes have diverged from nonplant actins at rates similar to those of nonplant actin genes from each other. If this is true, then it is logical to assume that these plant actin genes have also diverged from one another at this rate. The nucleotide replacement substitution between the three soybean actin genes, which ranged from 6–9%, is striking in light of these slow rates of replacement substitution. These data suggest that the three soybean actin genes had an ancient origin within the plant kingdom. Considering the twofold range of rates observed for animal, fungal and plant actin genes, we estimate a minimal divergence time of 330–660 MY for the three soybean genes.

A phenogram, using percentage replacement substitution data (Figure 3), illustrates the evolutionary relationships of actin genes from plant, fungal and animal sources. From this phenogram, it can be seen that soybean actin genes *SAc1*, *SAc3* and *SAc4* are more divergent (greater percentage of replacement substitution) from one another than is higher animal cytoplasmic actin from higher animal muscle actin.

Soybean encodes three classes of actin: The three soybean actin genes examined in this study are members of a multigene family, and hybridization studies have shown that six genes in this family fall into three classes, with a pair of genes in each class (HIGHTOWER and MEAGHER 1985). In Figure 2, the replacement substitution values for *SAc1*, *SAc3* and *SAc4*, representing each class, are given. These replacement substitution data on the soybean actin genes, ranging from 6–9%, as well as the hybridization data demonstrate the great divergence of gene family members. We have named these three classes as follows: kappa (K), containing *SAc1* and *SAc6* soybean actin genes; lambda (L), containing *SAc2* and *SAc4* soybean actin genes; and mu (M), containing *SAc3* and *SAc7* soybean actin genes.

Differential expression and evolution of actin genes within a multigene family: Members of actin gene families appear to be differentially expressed in those organisms for which actin gene expression has been examined (VANDEKERCKHOVE and WEBER 1978; TSANG, MAHBUBANI and WILLIAMS 1982; FYRBERG *et al.* 1983). The isoforms of nonplant actin do not appear to have specific functional differences in cellular processes, but do appear to be differentially expressed in striated muscle, smooth muscle and nonmuscle tissues (GUNNING *et al.* 1984).

Considering the variety of functions suggested for plant actins and the level of divergence between them, we would expect that the members of the soybean actin gene family are also differentially expressed. Preliminary studies with RNA dot blots and Northern blots suggest that there is organ-specific expres-

sion for two of the five soybean actin genes examined (HIGHTOWER and MEAGHER 1985). The plant organs examined, root, shoot and hypocotyl, contain a variety of cell types including cortical, vascular and epidermal tissues. The hypothesis that the diverse actin genes in soybean are differentially expressed at the tissue level is currently being investigated.

Differential levels of actin protein have been observed in plant cells and tissues. Cells in the vascular tissue of conifer roots that display cytoplasmic streaming contain greater amounts of actin than do adjacent tissues (PESACRETA *et al.* 1982). METCALF *et al.* (1984) observed the highest frequency of microfilaments in soybean root tips and radicles, and they estimated that actin concentrations in root tip extracts were 15-fold higher than those of leaf and petiole. The organization and distribution of F-actin during the cell cycle of meristematic onion root tip cells has been examined by CLAYTON and LLOYD (1985). These studies indicated that actin is present during cell division in the cytokinetic phragmoplast and is codistributed with microtubules at all stages of cell plate formation. F-actin has also been observed in actively cytoplasmic streaming rootcap cells from wheat and soybean (M. HAWES, personal communication; this laboratory).

Fossils of vascular land plants appear in samples from the late Silurian period, approximately 400 MY ago (TAYLOR 1981). In the 50 MY that followed, the Devonian period, land plants with a complex vascular tissue resembling a stele or protostele became widespread. These plants contain many cell and tissue types present in higher plants, including xylem, phloem, meristem, stomata and epidermis. The divergence (330–660 MY) of actin genes suggested by our data might reflect selection pressures for differentially expressed actin genes in new tissues and cell types.

Muscle actin: VANDEKERCKHOVE and WEBER (1978) analyzed amino acid sequences of mammalian and avian actins and observed different isoforms of actin. Warm-blooded vertebrates contain six actins: two cytoplasmic-specific actins and four muscle-specific actins: skeletal and cardiac striated muscle, vascular and nonvascular smooth muscle actins. The muscle actins were found to be very similar to one another, whereas the cytoplasmic actins were similar to actins found in Dictyostelium, Physarum and Acanthamoeba (VANDEKERCKHOVE and WEBER 1980; NELLEN and GALLWITZ 1982). They concluded that an ancestral muscle actin gene expressed in lower vertebrates was duplicated before or during early amphibian evolution, resulting in the evolution of a striated muscle and a smooth muscle actin gene. A proposed second duplication event could have given rise to cardiac and skeletal striated muscle actin genes, and vascular and nonvascular smooth muscle actin genes (VANDEKERCKHOVE, DECOUET and WEBER 1983).

Comparisons between human smooth:human cardiac, human smooth:rat skeletal and human cardiac:rat skeletal actin gene sequences indicate ~1% replacement substitution. This small value is compatible with the hypothesis that a more recent, second gene duplication event gave rise to these genes.

Oxytricha, a ciliated protozoan: The nucleotide sequence of the Oxytricha actin gene showed 29–33% replacement substitution with respect to all the

other actin genes examined. The hydropathy profiles of Oxytricha actin compared with the other actin sequences indicated less similarity (0.77–0.82) than any other pairwise comparisons (0.91–1.00). The Oxytricha actin is composed of 356 amino acids; 18 amino acids between residues 68–85 are absent in comparison to the other actin gene sequences examined. KAINE and SPEAR (1982) suggested that this actin had possibly become highly specialized in function. Ciliated protozoans have an elaborate, highly ordered anatomical structure, and it is plausible that proteins comprising this structure may have been uniquely modified.

Interactions of actin with other proteins: Comparisons of hydropathy profiles illustrate the highly conserved nature of actin from plant and nonplant sources. The number and lengths of hydrophilic and hydrophobic domains are conserved, and most of the replacement substitutions between actins have resulted in conservative amino acid replacements. This conservation is thought to reflect strong selection pressures acting to preserve the tertiary protein structure, therefore maintaining essential sites for the interaction of actin binding proteins and for the assembly of actin monomers to form polymeric filaments (POLLARD 1984).

Approximately 60 actin binding proteins have been characterized in animals (POLLARD 1984), whereas no actin binding proteins other than myosin have been identified from plants (LLOYD 1983). Cytoplasmic microfilaments occur in plant cells as single filaments or in bundles. Little is known about the development of microfilament arrays in plant cells or about the factors that regulate bundle formation. Based on ultrastructural studies, cytoskeletal components in plant and animal cells have similar features except for the apparent absence of intermediate filaments in plant cells (TIWARI *et al.* 1984). However, recently, a monoclonal antibody raised against an intermediate filament antigen was shown to cross-react with an antigen in higher plants (DAWSON, HULME and LLOYD 1985).

The amino acid conservation between plant and nonplant actins suggests that some of the actin binding proteins discovered thus far, or analogues of them, may be found in plant cells. Additional evidence to support this proposal is the conservation of amino acid residues in plant and nonplant actins that are thought to perform a functional role in actin structure and dynamics (LEAVIS and GERGELY 1984). Particular sulfhydryl groups on actins are required for ATP binding, for polymerization and for interaction with myosin. Cysteine residues 287 and 376, the latter thought to be directly involved in myosin binding, are conserved in all actins. Cysteine residues 12, 219 and 259 are conserved in most of the actins examined. Tyrosine residue 71, thought to be involved in the polymerization process (ELZINGA and COLLINS 1972), is conserved in all actins examined. LU and SZILAGYI (1981) have measured reductions in the reactivities of lysine residues 63, 70, 115, 286 and 338 on the protein surface after polymerization has occurred. Lysine residue 240 appears to be involved in the interaction of actin with tropomyosin (EL-SALEH *et al.* 1984). With the exception of residue 70 in yeast, all of these residues are conserved in all actins examined. Therefore, it seems plausible that some of

the proteins associated with regulation of actin polymerization in animal cells, or analogues of them, may be identified in plant cells.

Conclusions about the molecular evolution of actin

1. Soybean actin gene sequences appear to have diverged from nonplant actin gene sequences at overall rates similar to those rates of divergence of most nonplant actins from each other (SHAH, HIGHTOWER and MEAGHER 1983).

2. Soybean contains at least three distinct sequence classes of actin (HIGHTOWER and MEAGHER 1985). The three classes are more divergent in nucleotide sequence from one another than is higher animal cytoplasmic actin from muscle actin. Their divergence is ancient, probably greater than 300 MY ago and, perhaps, coinciding with the emergence and subsequent differentiation of vascular plants.

3. Amino acid substitutions are randomly distributed in actin proteins from plant and nonplant sources examined as a whole, and most of the amino acid changes between actins have resulted in conservative amino acid replacements.

4. The evolutionarily conserved nature of actin from plant and nonplant sources, as indicated by hydropathy profiles, conservative amino acid replacements and amino acid homology, suggests that some of the actin binding proteins that have been characterized in animal cells, or analogues of them, may be found in plant cells.

5. The physiological roles of actin in plant and animal cells suggest that similarities and differences may be found in actins from these sources. Such similarities and differences may reflect shared and specialized functions, respectively, of actin in these cells. Additional data on plant actins are essential to broaden the understanding of the molecular evolution of actin genes and proteins.

We would like to thank M. MCLEAN, T. MCKNIGHT, L. KEDES and H. ERBA for supplying unpublished data. We would also like to thank J. HIGHTOWER for assistance in analyzing hydropathy data. We have appreciated editorial comments from G. GALAU, J. STROMMER, V. BAIRD, L. PEARSON, B. BERMINGHAM and M. MCLEAN. We would like to acknowledge the expertise provided by K. RICE in preparing the DIV package of programs, and J. ARNOLD for advice on constructing the phenogram. This research was supported by grants from the National Science Foundation and the United States Department of Agriculture. R.C.H. was supported by a National Institutes of Health predoctoral genetics training grant. This work serves as partial fulfillment of the requirements for a Ph.D. degree in genetics by R.C.H.

LITERATURE CITED

AN, G. and J. D. FRIESEN, 1980 The nucleotide sequence of TufB and four nearby tRNA structural genes of *Escherichia coli*. Gene **12:** 33–39.

BERGSMA, D. J., K. S. CHANG and R. J. SCHWARTZ, 1985 Novel chicken actin gene: third cytoplasmic isoform. Mol. Cell. Biol. **5:** 1151–1162.

BLATT, M. R., N. K. WESSELLS and W. R. BRIGGS, 1980 Actin and cortical fiber reticulation in the siphonaceous alga *Vaucheria sessilis.* Planta **147:** 363–375.

BROWN, G. G., J. S. LEE, N. BRISSON and D. P. S. VERMA, 1984 The evolution of a plant globin gene family. J. Mol. Evol. **21:** 19–32.

CLAYTON, L. and C. W. LLOYD, 1985 Actin organization during the cell cycle in meristematic plant cells. Exp. Cell Res. **156:** 231–238.

COOPER, A. D. and W. R. CRAIN, JR., 1982 Complete nucleotide sequence of a sea urchin actin gene. Nucleic Acids Res. **10:** 4081–4092.

DAVIDSON, E. H., T. L. THOMAS, R. H. SCHELLER and R. J. BRITTEN, 1982 The sea urchin actin genes, and a speculation on the evolutionary significance of small gene families. pp. 177–191. In: *Genome Evolution.* Academic Press, New York.

DAWSON, P. J., J. S. HULME and C. W. LLOYD, 1985 Monoclonal antibody to intermediate filament antigen cross-reacts with higher plant cells. J. Cell Biol. **100:** 1793–1798.

EFSTRATIADIS, A., J. W. POSAKONY, T. MANIATIS, R. M. LAWN, C. O'CONNELL, R. A. SPRITZ, J. D. DERIEL, B. G. FORGET, S. M. WEISSMAN, J. L. SLIGHTOM, A. E. BLECHL, O. SMITHIES, F. E. BARALLE, C. C. SHOULDERS and N. J. PROUDFOOT, 1980 The structure and evolution of the human beta-globin gene family. Cell **21:** 653–668.

EL-SALEH, S. C., R. THIERET, P. JOHNSON and J. D. POTTER, 1984 Modification of lys-237 on actin by 2-4-pentanedione. J. Biol. Chem. **259:** 11014–11021.

ELZINGA, M. and J. H. COLLINS, 1972 The amino acid sequence of rabbit skeletal muscle actin. Cold Spring Harbor Symp. Quant. Biol. **1:** 37.

ENGEL, J. N., P. W. GUNNING and L. H. KEDES, 1981 Isolation and characterization of human actin genes. Proc. Natl. Acad. Sci. USA **78:** 4674–4678.

FILES, J. G., S. CARR and D. HIRSH, 1983 The actin gene family of *Caenorhabditis elegans.* J. Mol. Biol. **164:** 355–375.

FORNWALD, J. A., G. KUNCIO, I. PENG and C. P. ORDAHL, 1982 The complete nucleotide sequence of the chicken alpha-actin gene and its evolutionary relationship to the actin gene family. Nucleic Acids Res. **10:** 3861–3876.

FYRBERG, E. A., B. J. BOND, N. D. HERSHEY, K. S. MIXTER and N. DAVIDSON, 1981 The actin genes of *Drosophila*: protein coding regions are highly conserved but intron positions are not. Cell **24:** 107–116.

FYRBERG, E. A., J. W. MAHAFFEY, B. J. BOND and N. DAVIDSON, 1983 Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. Cell **33:** 115–123.

GUNNING, P., P. PONTE, L. KEDES, R. J. HICKEY and A. I. SKOULTCHI, 1984 Expression of human cardiac actin in mouse L cells: a sarcomeric actin associates with a nonmuscle cytoskeleton. Cell **36:** 709–715.

HAMADA, H., M. G. PETRINO and T. KAKUNAGA, 1982 Molecular structure and evolutionary origin of human cardiac muscle actin gene. Proc. Natl. Acad. Sci. USA **79:** 5901–5905.

HIGASHI-FUJIME, S., 1980 Active movement *in vitro* of bundles of microfilaments isolated from *Nitella* cell. J. Cell. Biol. **87:** 569–578.

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

ILKER, R. A., R. W. BREIDENBACH and T. M. MURPHY, 1979 Partial purification of actin from wheat germ. Phytochemistry (Oxf.) **18:** 1781–1783.

JACKSON, W. T. and B. G. DOYLE, 1977 Characterization of actin from root tips of *Phaseolus vulgaris.* J. Cell. Biol. **75:** 268a.

KAINE, B. P. and B. B. SPEAR, 1982 Nucleotide sequence of a macro nuclear gene for actin in *Oxytricha fallax.* Nature **295:** 430–432.

KLEIN, K., G. WAGNER and M. R. BLATT, 1980 Heavy-meromyosin decoration of microfilaments from *Mougeotia* protoplasts. Planta **150:** 354–356.

KYTE, J. and R. F. DOOLITTLE, 1982 A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:** 105–132.

LEAVIS, P. C. and J. GERGELY, 1984 Thin filament proteins and thin filament-linked regulation of vertebrate muscle contraction. CRC Critical Rev. Biochem. **16:** 235–305.

LEWIN, B., 1983 Genes. pp. 350–352. John Wiley & Sons, New York.

LLOYD, C. W., 1983 *The Cytoskeleton in Plant Growth and Development.* pp. 3–29. Academic Press, London.

LU, R. C. and L. SZILAGYI, 1981 Change of reactivity of lysine residues upon actin polymerization. Biochemistry **20:** 5914–5919.

MARGULIS, L. and K. V. SCHWARTZ, 1976 *Five Kingdoms.* pp. 17 and 161. W. H. Freeman, San Francisco.

MCKEOWN, M. and R. A. FIRTEL, 1981 Differential expression and 5′ end mapping of actin genes in *Dictyostelium.* Cell **24:** 799–807.

METCALF III, T. N., L. J. SZABO, K. R. SCHUBERT and J. L. WANG, 1980 Immunochemical identification of an actin-like protein from soybean seedlings. Nature **285:** 171–172.

METCALF III, T. N., L. J. SZABO, K. R. SCHUBERT and J. L. WANG, 1984 Ultrastructural and immunochemical analyses of the distribution of microfilaments in seedlings and plants of *Glycine max.* Protoplasma **120:** 91–99.

NELLEN, W. and D. GALLWITZ, 1982 Actin genes and actin messenger RNA in *Acanthamoeba castellanii.* Nucleotide sequence of the split gene I. J. Mol. Biol. **159:** 1–18.

NG, R. and J. ABELSON, 1980 Isolation and sequence of the gene for actin in *Saccharomyces cerevisiae.* Proc. Natl. Acad. Sci. USA **77:** 3912–3916.

NUDEL, U., R. ZAKUT, M. SHANI, S. NEUMAN, Z. LEVY and D. YAFFE, 1983 The nucleotide sequence of the rat cytoplasmic beta-actin gene. Nucleic Acids Res. **11:** 1759–1771.

PERLER, F., A. EFSTRATIADIS, P. LOMEDICO, W. GILBERT, R. KOLODNER and J. DODGSON, 1980 The evolution of genes—the chicken preproinsulin gene. Cell **20:** 555–566.

PESACRETA, T. C., W. W. CARLEY, W. W. WEBB and M. V. PARTHASARTHY, 1982 F-actin in conifer roots. Proc. Natl. Acad. Sci. USA **79:** 2898–2901.

POLLARD, T. D., 1984 Actin-binding protein evolution. Nature **312:** 403.

PONTE, P., P. GUNNING, H. BLAU and L. KEDES, 1983 Human actin genes are single copy for an alpha-skeletal and alpha-cardiac actin but multicopy for beta- and gamma-cytoskeletal genes: 3′ untranslated regions are isotype specific but conserved in evolution. Mol. Cell. Biol. **3:** 1781–1791.

PONTE, P., S-Y. NG, J. ENGEL, P. GUNNING and L. KEDES, 1984 Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human beta-actin cDNA. Nucleic Acids Res. **12:** 1687–1696.

SANCHEZ, F., S. L. TOBIN, U. RDEST, E. ZULAUF and B. J. MCCARTHY, 1983 Two *Drosophila* actin genes in detail: gene structure, protein structure and transcription during development. J. Mol. Biol. **163:** 533–551.

SAS INSTITUTE, 1982 *SAS User's Guide: Basics, 1982 Edition.* pp. 501–512. SAS Institute, Cary, North Carolina.

SCHLIWA, M., 1981 Proteins associated with cytoplasmic actin. Cell **25:** 587–590.

SCHULER, M. A., P. MCOSKAR and E. B. KELLER, 1983 DNA sequence of two linked actin genes of sea urchin. Mol. Cell. Biol. **3:** 448–456.

SHAH, D. M., R. C. HIGHTOWER and R. B. MEAGHER, 1982 Complete nucleotide sequence of a soybean actin gene. Proc. Natl. Acad. Sci. USA **79:** 1022–1026.

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. Genet. **2:** 111–126.

SOKAL, R. R. and P. H. A. SNEATH, 1963 *Principles of Numerical Taxonomy.* p. 189. W. H. Freeman, San Francisco.

STOSSEL, T. P., 1984 Contribution of actin to the structure of the cytoplasmic matrix. J. Cell. Biol. **99:** 15s–21s.

TAYLOR, T. N., 1981 The evolution of vascular plants. pp. 77–91. In: *Paleobotany: An Introduction to Fossil Plant Biology.* McGraw-Hill, New York.

TIWARI, S. C., S. M. WICK, R. E. WILLIAMSON and B. E. S. GUNNING, 1984 Cytoskeleton and integration of cellular function in cells of higher plants. J. Cell. Biol. **99:** 63s–69s.

TSANG, A. S., H. MAHBUBANI and J. G. WILLIAMS, 1982 Cell-type-specific actin mRNA populations in *Dictyostelium discoideum.* Cell **31:** 375–382.

UEYAMA, H., H. HAMADA, N. BATTULA and T. KAKUNAGA, 1984 Structure of a human smooth muscle actin gene (aortic type) with a unique intron site. Mol. Cell. Biol. **4:** 1073–1078.

VAHEY, M. and S. P. SCORDILIS, 1980 Contractile proteins from the tomato. Can. J. Bot. **58:** 797–801.

VANDEKERCKHOVE, J., H-G. DECOUET and K. WEBER, 1983 Molecular evolution of muscle-specific actins: a protein-chemical analysis. pp. 241–248. In: *Actin Structure and Function in Muscle and Non-muscle Cells.* Academic Press, Australia.

VANDEKERCKHOVE, J. and K. WEBER, 1978 At least six different actins are expressed in a higher mammal: an analysis based on the amino acid sequence of the amino-terminal tryptic peptide. J. Mol. Biol. **126:** 783–802.

VANDEKERCKHOVE, J. and K. WEBER, 1980 Vegetative *Dictyostelium* cells containing seventeen actin genes express a single major actin. Nature **284:** 475–477.

WILLIAMSON, R. E., 1975 Cytoplasmic streaming in *Chara*: a cell model activated by ATP and inhibited by cytochalasin B. J. Cell. Sci. **17:** 655–668.

WILLIAMSON, R. E., 1980 Actin in motile and other processes in plant cells. Can. J. Bot. **58:** 766–772.

WILSON, A. C., S. S. CARLSON and T. J. WHITE, 1977 Biochemical evolution. Annu. Rev. Biochem. **46:** 573–639.

ZAKUT, R., M. SHANI, D. GIVOL, S. NEUMAN, D. YAFFE and U. NUDEL, 1982 Nucleotide sequence of the rat skeletal muscle actin gene. Nature **298:** 857–859.

Communicating editor: C. C. LAURIE-AHLBERG