J Mol Evol (1986) 23:11–22

Journal of Molecular Evolution

© Springer-Verlag New York Inc. 1986

Comparison of Three Actin-Coding Sequences in the Mouse; Evolutionary Relationships Between the Actin Genes of Warm-Blooded Vertebrates

Serge Alonso, Adrian Minty,* Yves Bourlet,† and Margaret Buckingham

Department of Molecular Biology, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France

Summary. We have determined the sequences of three recombinant cDNAs complementary to different mouse actin mRNAs that contain more than 90% of the coding sequences and complete or partial 3′ untranslated regions (3′UTRs): pAM 91, complementary to the actin mRNA expressed in adult skeletal muscle (α_{sk} actin); pAF 81, complementary to an actin mRNA that is accumulated in fetal skeletal muscle and is the major transcript in adult cardiac muscle (α_c actin); and pAL 41, identified as complementary to a β nonmuscle actin mRNA on the basis of its 3′UTR sequence.

As in other species, the protein sequences of these isoforms are highly (>93%) conserved, but the three mRNAs show significant divergence (13.8–16.5%) at silent nucleotide positions in their coding regions. A nucleotide region located toward the 5′ end shows significantly less divergence (5.6–8.7%) among the three mouse actin mRNAs; a second region, near the 3′ end, also shows less divergence (6.9%), in this case between the mouse β and α_{sk} actin mRNAs. We propose that recombinational events between actin sequences may have homogenized these regions. Such events distort the calculated evolutionary distances between sequences within a species.

Codon usage in the three actin mRNAs is clearly different, and indicates that there is no strict relation between the tissue type, and hence the tRNA precursor pool, and codon usage in these and other muscle mRNAs examined. Analysis of codon usage in these coding sequences in different vertebrate species indicates two tendencies: increases in bias toward the use of G and C in the third codon position in paralogous comparisons (in the order $\alpha_c < \beta < \alpha_{sk}$), and in orthologous comparisons (in the order chicken < rodent < man).

Comparison of actin-coding sequences between species was carried out using the Perler method of analysis. As one moves backward in time, changes at silent sites first accumulate rapidly, then begin to saturate after −(30–40) million years (MY), and actually decrease between −400 and −500 MY. Replacements or silent substitutions therefore cannot be used as evolutionary clocks for these sequences over long periods. Other phenomena, such as gene conversion or isochore compartmentalization, probably distort the estimated divergence time.

Key words: Actin-coding regions — Sequence divergence — Conversion — Codon usage — Evolution

Introduction

Actin is a protein found in all eukaryotic cells so far examined, and its amino acid sequence is highly (>90%) conserved among species. This abundant structural protein is important in the maintenance

* *Present address:* INSERM V129, Unité de Recherches en Genétique et Pathologie Moléculaire, 24, rue du Paubowig, St. Jacques, 75674 Paris Cedex 14, France

† *Present address:* Institut d'Embryologie, 49bis avenue de la belle Gabrielle, 94130 Nogent sur Marne, France

Offprint requests to: S. Alonso

of cell shape and in cell motility. In nonmuscle cells polymerized actin is the main component of the microfilaments of the cytoskeleton (for a review, see Pollard and Weihing 1974), while in muscle it is a major contractile protein of the sarcomere (for a review, see Bagshaw 1982).

With the exception of the single actin gene in yeast (Gallwitz and Seidel 1980; Ng and Abelson 1980), in all organisms so far examined the actins constitute a family of highly conserved proteins encoded by several differentially regulated genes (Buckingham and Minty 1983). In warm-blooded vertebrates at least six actin variants have been characterized: two from smooth muscles (α_{sm} and γ_{sm}), two from striated muscles (α_c, the major isoform of adult cardiac muscle, and α_{sk}, that of adult skeletal muscle), and two from nonmuscle tissues (β and γ) (Vandekerckhove and Weber 1979). A third cytoplasmic actin has recently been described in chicken (Bergsma et al. 1985). The tissue specificity of muscle isoforms might suggest that their structural conservation is related to the function of each isoform in its particular cell type. However, some observations argue against this point of view. First, the actins present in the muscle tissues of invertebrates resemble the vertebrate cytoplasmic actins rather than the vertebrate muscle actins (e.g., see Vandekerckhove et al. 1983; Vandekerckhove and Weber 1984); the actin isoform characteristic of *Drosophila* flight muscle, for example, has a vertebrate-cytoplasmic-type actin sequence (Fyrberg et al. 1981). Second, vertebrate actin isoforms seem to be to some extent interchangeable. Thus skeletal and cardiac muscle actins are coexpressed in the same tissue during striated muscle development (Minty et al. 1982; Mayer et al. 1984; Vandekerckhove et al. 1986), and cardiac actin has been shown to participate in the formation of cytoskeletal structures on the introduction of a human cardiac actin gene into a nonmuscle mouse L cell (Gunning et al. 1984a). An alternative, extreme point of view is that the different actin isoforms confer no distinct functional advantage, but have evolved because of a regulatory requirement for multiple actin genes (Davidson and Britten 1973). This seems inherently less likely. Nevertheless, it is clear from the preceding discussion that if the development of the distinct muscle-type sequences seen in warm-blooded vertebrates represents an optimization of muscle-actin function, considerable functional flexibility exists between muscle- and nonmuscle-type isoforms.

The appearance during evolution of muscle-type actin sequences has been traced at the protein level in primitive chordates and also in amphibia (Vandekerckhove and Weber 1984). A muscle-type actin is found in the lamprey that presumably arose from a nonmuscle-type ancestral actin gene by duplication and sequence divergence during early chordate evolution. Subsequently, similar events led to the appearance of two distinct muscle sequences, one smooth type and one striated type muscle (salamander), and then to two striated muscle actins, one cardiac-like and one skeletal-like muscle (*Xenopus*), during early amphibian evolution, 300–450 million years (MY) ago.

In most mammalian genomes, a large number (more than 20) of actin-related sequences can be detected by Southern blot analysis or by cloning of genomic DNA (Engel et al. 1981; Humphries et al. 1981; Minty et al. 1983). Since the six actin isoforms identified in mammals are probably encoded by single genes (see Minty et al. 1983; Ueyama et al. 1984), many of the actin-related genomic sequences may represent pseudogenes (Moos and Gallwitz 1982; Minty et al. 1983). The majority of these sequences are homologous to cytoplasmic β or γ actin mRNAs (Ponte et al. 1983) and are dispersed in the mammalian genome. This is also the case for the structural genes encoding the actin isoforms: No genetic linkage is seen between α_{sk}, α_c, and β actin genes in the mouse (Czosnek et al. 1983; Minty et al. 1983; Robert et al. 1985), nor between α_{sk} and α_c genes in humans (Gunning et al. 1984b).

Nucleotide sequences give more information on the process of molecular evolution than do protein sequences, especially when the protein sequences are so highly conserved as those of the actins. In this paper we compare three actin-coding sequences in the mouse corresponding to α_{sk}, α_c, and β actin gene transcripts, in terms of sequence divergence and codon usage. This comparison is then extended to an analysis of these actin-coding sequences in other mammals and birds. Perler's method (Perler et al. 1980) of obtaining the correct percentage divergence between two nucleotide sequences plotted against their time of divergence (estimated mainly from the fossil record) has been applied to these interspecies comparisons. The results indicate that for this multigene family the divergence of actin-coding sequences can be used as an evolutionary clock over limited periods only.

Experimental Procedures

Construction and Nucleic-Acid Sequencing of cDNA Plasmids. The construction of recombinant cDNA plasmids has been described by Minty et al. (1981). Restriction fragments were end-labeled at their 5′ or 3′ extremities. For 5′ labeling, restriction fragments were labeled with T_4 polynucleotide kinase (Bethesda Research Laboratories) using the exchange reaction (see Minty et al. 1981). For 3′ labeling, protruding 3′ ends were labeled with terminal deoxynucleotidyltransferase (Boehringer Mannheim) by the addition of one [α-^{32}P]ddATP residue (Åmersham; 3000 Ci/

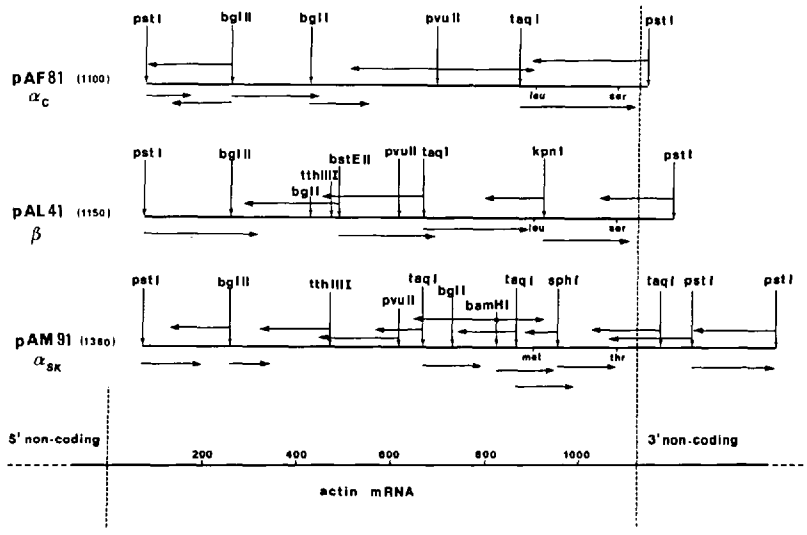


Fig. 1. Restriction maps of plasmids pAF 81, pAL 41, and pAM 91. Maps have been aligned. Horizontal arrows indicate regions sequenced. Arrows above the line indicate sequencing of the mRNA strand; those below the line indicate sequencing of the cDNA strand. Numbers in parentheses correspond to the size in base pairs of each insertion. We indicate also the positions of the codons encoding the two key residues 298 and 357 in each cDNA. Vertical arrows mark restriction sites. The dashed vertical line indicates the end of the coding sequence

mmol) (see Robert et al. 1982) and recessed 3′ ends were labeled with DNA polymerase I, Klenow fragment (Boehringer Mannheim) according to Challberg and Englund (1980). Labeled restriction fragments were then strand-separated in 5% polyacrylamide gels [5% polyacrylamide, 0.08% methylene bisacrylamide, 50 mM Tris-borate, pH 8.3, 1 mM ethylenediaminetetraacetate (EDTA)], or they were digested with a second restriction enzyme and the two labeled fragments then separated in 5% polyacrylamide (5% polyacrylamide, 0.17% methylene bisacrylamide, 50 mM Tris-borate, pH 8.3, 1 mM EDTA) gels. After electroelution of DNA from the gels, the fragments were sequenced by the chemical degradation method of Maxam and Gilbert (1980) using 0.35 mm thick 6%, 8%, or 20% polyacrylamide gels (Sanger and Coulson 1978).

Computer Analyses. Nucleotide sequences were analyzed using a self-serve sequence-analysis system implemented on a Data General MV 8000 computer at the Pasteur Institute. Computer programs that permitted the comparison of nucleotide sequences, searches for codon usage, and calculation of total G + C content were adapted by Claverie (1984) from programs provided by Staden (Staden and McLachlan 1982).

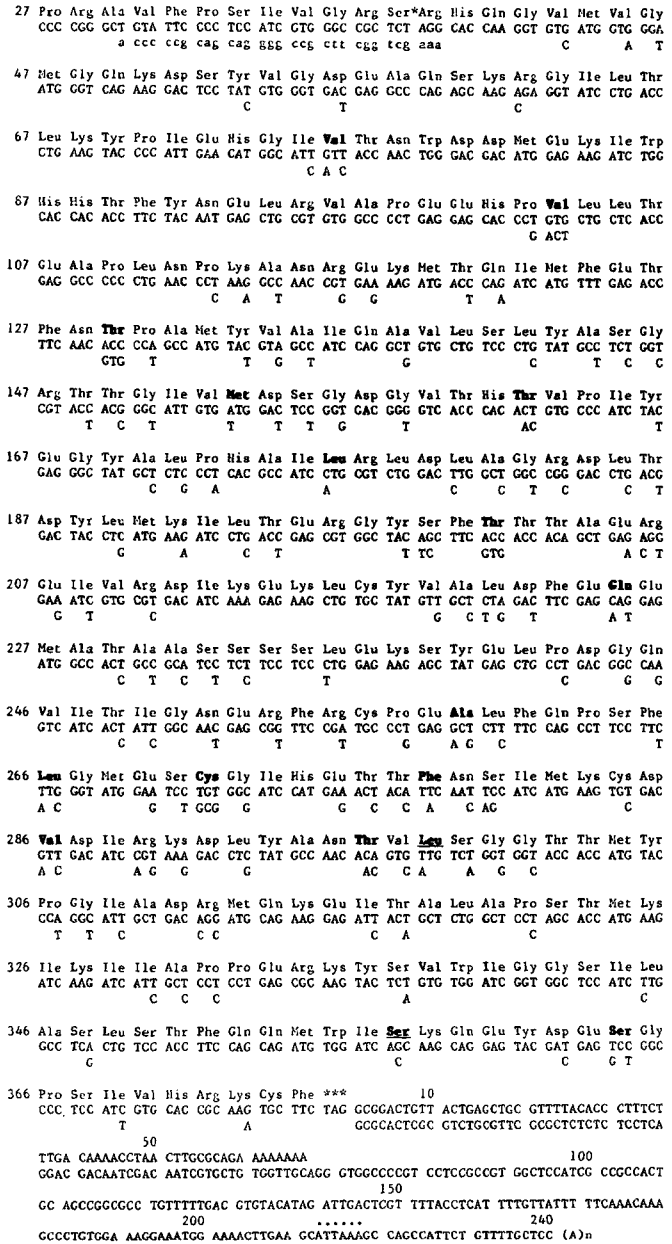
Results and Discussion

I. Isolation and Characterization of Three Mouse Actin cDNA Plasmids

Recombinant cDNA plasmids pAM 91 and pAF 81 were cloned from poly(A)+ RNA isolated from skeletal muscle of 10-day-old mice (Minty et al. 1981). Partial nucleotide sequences, together with results of RNA and DNA hybridization experiments done to characterize the tissue expression of the homologous mRNA and to demonstrate that it is the product of a single gene, led to the conclusion that pAM 91 is complementary to the actin mRNA expressed in adult skeletal muscle (α_{sk}) (Minty et al. 1981) and that pAF 81 is complementary to a fetal skeletal muscle actin mRNA that is indistinguishable from that expressed in adult cardiac tissue (α_c) (Minty et al. 1982). Plasmid pAL 41 was isolated from a mouse lymphocyte cDNA library (Kvist et al. 1981) by cross hybridization with the pAM 91 probe. The restriction maps of these three cDNAs and the sequencing strategy adopted are presented in Fig. 1.

The sequence of pAM 91 encodes the same protein sequence as that published for bovine skeletal muscle actin (Vandekerckhove and Weber 1979). At the extreme 5′ end of pAM 91 there is a sequence inversion of the region covering amino acids 18 to 27, followed by a deletion of amino acids 28–39, probably a cloning artifact (Fields and Winter 1981; Volckaert et al. 1981). Plasmid pAM 91 contains the complete 240-nucleotide 3′ untranslated region (3′UTR) of the α_{sk} actin mRNA; the polyadenylation signal ATTAAA is located 28 base pairs upstream from the poly(dA) tail (Fig. 2A). Comparison of this sequence with that of the rat α_{sk} 3′UTR (Zakut et al. 1982) demonstrates only seven substitutions and two point deletions in the rat sequence (which is 238 nucleotides long). The 3′UTR sequences are therefore as conserved (96.2%) as the coding sequences (96.3%) are between these two species. Significantly more sequence divergence is seen between the UTRs of mammals and birds (Ordahl and Cooper 1983).

Plasmid pAF 81 contains the coding sequence of the α_c actin mRNA together with the first 14 nucleotides of the 3′UTR (Fig. 2B), which matches (at 12 of 14 nucleotides) the sequence at the beginning of the human α_c actin 3′UTR (Hamada et al. 1982). This short sequence is missing in the rat, where the α_c actin 3′UTR is 163 nucleotides long (Mayer et al. 1984). The sequence of pAF 81 encodes a protein identical to the published sequence of bovine cardiac muscle actin (Vandekerckhove and Weber 1979) except for one amino acid replacement, a substi-


A
B

tution of histidine for arginine at position 370 (see Minty et al. 1982).

The sequence of pAL 41 (Fig. 2) encodes a cytoplasmic actin (β or γ) almost identical to the bovine cytoplasmic actins (Vandekerckhove and Weber 1979). It is probably a β actin, since the first 67 nucleotides of the gene's 3'UTR are identical to those from the beginning of the rat β actin 3'UTR (Nudel et al. 1983) and it has been shown that the 3'UTRs of β and γ actin mRNAs do not cross-hybridize in chicken (Cleveland et al. 1980) and human (Ponte et al. 1983). That the 3'UTR of pAL 41 is characteristic of β actin mRNA strongly suggests that pAL 41 encodes a β actin. The coding sequence of pAL 41 shows one amino acid replacement relative to the published sequence of bovine β cytoplasmic actin (Vandekerckhove and Weber 1979): Proline is substituted for serine at position 38. As in the case of the substitution between pAF 81 and the bovine cardiac actin sequence (position 370, his → arg), this nonconservative change is not found in any of the actin protein sequences established for a wide range of different species; this residue is invariant in these sequences. We therefore suspect that an enzymatic error, probably by reverse transcriptase, during the construction of the plasmids caused this change (see Minty et al. 1982).

II. Nucleotide-Sequence Comparisons Between pAM 91, pAF 81, and pAL 41

a. Overall Sequence Divergence. We have calculated the degree of divergence between the coding sequences contained in pAM 91, pAF 81, and pAL 41 as the total number of substitutions over the full length compared (Table 1). Because of the sequence inversion at the extreme 5' end in pAM 91, coding sequences were compared from the histidine codon at position 40 to the stop codon, a region representing some 90% (1011 of 1125 nucleotides) of the whole coding sequence. This comparison therefore excludes the amino-terminal coding region, where several of the few amino acid substitutions between isoforms are located (Vandekerckhove and Weber 1979). However, we verified that deletion of the first 39 codons of the complete actin gene sequences did not influence significantly the overall sequence divergence.

Table 1. Percentage nucleotide divergence between mouse cDNAS

Actin-coding regions	Overall nucleotide divergence[a] (%)	Local nucleotide divergence at the 5' end[b] (%)
α_c/α_{sk}	13.8 (13.5)	5.6
β/α_{sk}	14.2 (11.9)	5.6
β/α_c	16.5 (14.2)	8.7

[a] Number in brackets represents the percentage of silent substitutions

[b] Corresponds to a region 162 nucleotides long between codons 40 (histidine) and 93 (glutamate)

Although the actin protein sequences are highly conserved, their mRNAs vary significantly due to the presence of numerous silent substitutions leading to synonymous codon changes. Between pAM 91 and pAF 81, there are 140 substitutions, of which only 4 are replacement substitutions; these result in three amino acid exchanges (positions 298, 357, and 370). The least diverged nucleotide sequences are those of the two sarcomeric actin mRNAs (Table 1). The percentage divergences increase to 14.2% and 16.5% when we compare α_{sk} and α_c mRNAs, respectively, with nonmuscle β actin mRNA. It is striking that the very similar proteins α_c and α_{sk} actin (4 differences out of 375 amino acids) are encoded by mRNAs whose sequences are relatively diverged compared with that of nonmuscle β actin. Cardiac actin mRNA, which encodes a protein that is more similar to β actin (23 changes) than is α_{sk} actin (25 changes), is more divergent from β actin at the nucleotide level.

Hybridization experiments had already suggested that the nucleotide sequences of α_{sk} and β actin mRNAs are closer to each other than are α_c and β actin in RNA sequences. On Southern blots of mouse genomic DNA, under conditions of high-stringency washing (70°C, 0.1 × SSC), most of the mouse cytoplasmic-related sequences are detected by hybridization with pAM 91 or pAL 41, but not with pAF

Fig. 2. Nucleotide sequence of pAL 41 and comparisons with those of pAM 91 (A) and pAF 81 (B). The restriction fragments indicated in Fig. 1 were end-labeled and sequenced by the chemical degradation method of Maxam and Gilbert (1980). The base sequence of pAL 41 and the amino acids that it encodes are shown. Amino acids in bold type are specific for nonmuscle actins (β or γ), and the two key positions (298 and 357) that distinguish α_c and α_{sk} actin sequences are underlined. Below the sequence of pAL 41 are shown the nucleotides in the sequences of pAM 91 (A) and pAF 81 (B) where they differ from that of pAL 41. Amino acids are numbered following the convention discussed by Vandekerckhove and Weber (1979); that is, the serine residue overlooked in the initial numbering is considered as residue 234a. Unexpected codons at positions 38 in pAL 41 and 370 in pAF 81 are labeled by asterisks. At the extreme 5' end of pAM 91 is a sequence inversion of the region from amino acids 18 to 27, followed by a deletion of amino acids 28–39. Amino acids indicated in the region of amino acids 18–27 correspond to reading the nucleotide sequence of the complementary strand in the reverse direction, and these codons are written using small letters. Partial or complete 3'UTRs have been aligned, although they are not homologous. The polyadenylation signal ATTAAA in pAM 91 (A) is indicated by dots above the line

81, which under these conditions hybridizes only to the cardiac and skeletal actin genes (Minty et al. 1983). Melting curves derived from Southern blot hybridizations with the different actin plasmids show that homologous hybrids melt at 70–71°C and heterologous hybrids at a temperature 11–13°C lower (Minty et al. 1983). A melting temperature difference of 1°C in these experiments thus represents a nucleotide-sequence divergence of 1.3%.

Analysis of actin proteins has led to the proposal that α_{sk} and α_c actin genes arose from a primordial sarcomeric actin gene by a gene-duplication event during amphibian evolution (Vandekerckhove and Weber 1984). If this were the case, one would predict that α_{sk} and α_c mRNA sequences would be similarly diverged from nonmuscle actin in terms of silent substitutions, and more similar to each other. However, the nucleotide-sequence divergence among the mouse actin mRNAs is not in keeping with such an evolutionary relationship; factors that distort the figures for overall sequence divergence are discussed in the following sections.

b. Regional Sequence Divergence. Comparisons between the three cDNAs show a region of greater sequence similarity in the 5′ part of the cDNAs, considered here to be from the histidine codon at position 40 (immediately after the inversion in pAM 91) to the glutamate codon at position 93. This region, probably located in the third exon (see Buckingham and Minty 1983), is much less diverged between the three sequences than the overall coding sequence is (Table 1). This regional sequence homology is highly significant, since the probabilities of finding this particular distribution at random are 1/1000 (when α_c/α_{sk} or β/α_{sk} are compared) and 3.5/1000 (when α_c/β are compared).

A second region including the last 130 nucleotides of the coding region between codons 331–374, and probably located in the last exon, is also less diverged (6.9%), in this case between mouse β and α_{sk} actin cDNAs. A region of lower sequence divergence (6.9%) has been noted at a similar position in the rat β and α_{sk} actin genes (for which the overall divergence is 15.8%) (Nudel et al. 1983).

Such regions of sequence similarity are probably maintained by a mechanism of sequence homogenization that can act on multigene families (Baltimore 1981) and in which recombination takes place between parts of the genes only, as in the cases, for example, of the human fetal $^{G}\gamma$ and $^{A}\gamma$ globin genes (Slightom et al. 1980) or of the two adult human α globin genes (Zimmer et al. 1980). The 3′ sequence region is also much less diverged between α_{sk} and β actin mRNAs in chicken (8.5%) (Fornwald et al. 1982; Kost et al. 1983), rat (6.9%) (Zakut et al. 1982; Nudel et al. 1983), and human (6%) (Hanauer et al. 1983; Ponte et al. 1984); the overall divergences in these species are 16.5%, 15.8%, and 14%, respectively. This is true in paralogous comparisons only: The sequence is not conserved between species. That sequence similarity is intraspecific is more in keeping with the occurrence of relatively recent gene-conversion-type events (Baltimore 1981) than with evolutionary conservation of the primordial sequence in this region for functional or structural reasons. In protein sequences corresponding to the 3′ and 5′ regions of the actin-coding sequences where gene conversion is postulated to have taken place, amino acid residues characteristic of nonmuscle or striated muscle actins are present at positions 76, 357, and 364. In each case the amino acid change is a neutral substitution resulting from a single base change. We suggest that the selective pressure on these sequences has been sufficiently strong to result in the selection of back mutations at these positions after the sequence-homogenization event.

Sequence homogenization may not have occurred directly between the three genes concerned. Given the number of actin-related sequences in most mammalian genomes (Minty et al. 1983), recombination events of this kind in the mouse or rat may have taken place via another, actin-related genomic sequence. If sequence homogenization is restricted to the coding sequences, this might suggest a cDNA-mediated gene-conversion mechanism operating, for example, via an intronless processed pseudogene present in the genome (Moos and Gallwitz 1982), or via a free cDNA template, as has been proposed in the case of the U3 small-nuclear-RNA genes in mammals (Bernstein et al. 1983).

III. Comparison of Codon Usage Between Actin Genes in Different Species

We have calculated the percentage G + C content and the (G + C)/(A + T) bias in the third codon positions in the coding sequences of several actin genes and cDNAs (Table 2). The ratio (G + C)/(A + T) permits one to gain an immediate appreciation of the nucleotide bias toward G or C at the third positions of codons (each letter in this case represents the number of codons ending with the corresponding base) (see also Hanauer et al. 1983). Two tendencies can be distinguished (cf. Nudel et al. 1983). First, Table 2 shows that there is an increase in bias in paralogous comparisons, bias $\alpha_c <$ bias $\beta <$ bias α_{sk}, that is conserved in chicken, in rat, in mouse, and, most strikingly, in human. The rat and mouse biases are the same for β (2.7) and α_{sk} (3.4) mRNAs, and the partial (95-codon) nucleotide sequence of rat α_c actin cDNA (Mayer et al. 1984) shows a 50.5% G+ C content and a bias of 1.7, which is in agreement with the bias of 1.4

Table 2. Codon usage in different actin and $MLC1_F/MLC3_F$ mRNAs

	Actin									Myosin		
	α_c			β			α_{sk}			$MLC1_F/MLC3_F$		
	G + C	Bias	Ref.	G + C	Bias	Ref.	G + C	Bias	Ref.	G + C	Bias	Ref.
Human	54	2	1	60.6	5.5	2	61.6	8.2	3	–	–	–
Rat	–	–	–	56.8	2.7	4	57.6	3.4	5	53	1.6	6
Mouse	51.3	1.4	–	55.7	2.7	–	56.4	3.4	–	51.7	1.5	7
Chicken	49.1	1.2	8	53.4	1.8	9	55.9	2.7	10	51.4	1.5	11

We calculated the total percentage of G + C and also the ratio (G + C)/(A + T), which measures the bias toward G or C in the third positions of codons. The following references were used for sequence data: 1, Hamada et al. (1982); 2, Ponte et al. (1984); 3, Hanauer et al. (1983); 4, Nudel et al. (1983); 5, Zakut et al. (1982); 6, Periasamy et al. (1984); 7, Robert et al. (1984); 8, Chang et al. (1985); 9, Kost et al. (1983); 10, Fornwald et al. (1982); 11, Nabeshima et al. (1984)

for the mouse α_c actin sequence. As indicated by the nucleotide divergence of the sequences in the mouse, the differences in G + C content and bias in any one species are greater between the α_c and β than between the β and α_{sk} nucleotide sequences. One specific example illustrates this point: in the mouse α_c-actin-coding sequence, codons GAA and GAG are used for Glu, whereas GAG is preferentially used in the β and α_{sk} sequences. Differences in codon usage between α_c, β, and α_{sk} actin genes are clearly not related to the amino acid compositions of the corresponding proteins; α_c and α_{sk} actin mRNAs, which have the lowest and highest G + C content, respectively, encode the most related proteins (only 4 amino acid differences in 375 positions, compared with 23/375 and 25/375 differences, respectively, from β actin). These observations on codon usage for the overall sequences also in fact apply to the similar sequence regions. Thus for the 5′ regions that are similar in the three mouse actin mRNAs (α_{sk}, α_c, and β), the frequency of G or C in the third codon position is similar for pAM 91 and pAL 41 and significantly lower for pAF 81.

Table 2 reveals a second tendency in bias, one related to species. The percentage G + C content for a given actin-coding sequence (e.g., α_{sk}) increases from chicken (55.9%) to rodent to human (61.6%). Examination of the G + C content of another muscle-coding sequence, namely the gene for the alkali myosin light chains ($MLC1_F/MLC3_F$) of adult fast skeletal muscle, demonstrates the same tendency (Table 2), suggesting that this may be a more general phenomenom.

Different factors may influence codon usage and be functionally significant at the level of mRNA translation, such as structural features of individual tRNA molecules and the nature of the codon–anticodon reaction itself, which depends on the relative proportions of G–C and A–T linkages. Different tissues may have different tRNA pools, and the relative abundance of an mRNA and its protein product in a tissue may also be related to codon usage. The three actin mRNAs examined here are all major mRNA species, coding for a major protein type in the tissues concerned. Cardiac and skeletal actin mRNAs are major components of the total mRNA in adult heart and skeletal muscle tissues, respectively, and the corresponding proteins accumulate in these tissues. During skeletal muscle development, however, the two mRNAs are coexpressed (Minty et al. 1982; Vandekerckhove et al. 1986). The converse situation has been described in developing (e.g., Mayer et al. 1984) and indeed in some adult mammalian hearts (Gunning et al. 1983; Vandekerckhove et al. 1986). It would thus appear that the same cell (Minty et al. 1982; Mayer et al. 1984) can express both mRNAs efficiently. Comparing codon usage of other genes expressed in cardiac or skeletal muscle (Table 3), we observe (see also Wain-Hobson et al. 1981; Hastings and Emerson 1983) that within a species synonymous-codon usage may follow similar patterns in different tissues or different patterns in the same tissue, and there is thus no evidence suggesting tissue-specific coadaptation of codon usage and tRNA pools.

Another aspect of the question of codon usage concerns the genomic organization of the sequences. Genes have been shown to be located in different isochores (Bernardi et al. 1985), composed of long stretches (several hundred kilobase pairs) of DNA of relatively uniform and distinct base composition. The total G + C contents of the exons of genes, including the mouse actin sequences discussed here, have been shown to be linearly related to those of the isochores in which they are embedded (Bernardi et al. 1985), suggesting that the isochore may condition the codon usage of genes in it. In the mouse, actin genes constitute a dispersed multigene family (Minty et al. 1983). When duplicated sequences are dispersed to different regions of the genome that either already differ in base composition or subsequently evolve toward different total G + C contents, silent substitutions in the genes concerned may follow different tendencies corresponding to those

Table 3. Codon usage in different genes (or cDNAs) expressed either in cardiac or skeletal muscle

	% G or C in the third position of codons	Bias	Number of codons examined	Ref.
Heart				
Rat α_c actin	64	1.8	94	1
Rat α MHC	86	6.1	427	2
Rat β MHC	85.5	6.0	425	2
Mouse α_c actin	59	1.4	350	
Mouse MHC	77	3.3	65	3
Chicken $MLC2A_F$	72	2.6	153	4
Chicken G3PDH	59	1.4	334	5
Skeletal muscle				
Rat $MLC1_F/MLC3_F$	60	1.6	198	6
Rat $MLC2_F$	70	2.3	169	7
Rat α_{sk} actin	77.5	3.4	375	8
Mouse α_{sk} actin	77	3.4	337	
Mouse MHC_F	77	3.4	166	9
Mouse $MLC1_F/MLC3_F$	62	1.5	198	10
Chicken $MLC1_F/MLC3_F$	60	1.5	198	11
Chicken α_{sk} actin	73	2.7	375	12

We calculated the percentage of G + C in the third positions of codons and the ratio (G + C)/(A + T), which can be used to represent the bias. References: 1, Mayer et al. (1984); 2, Mahdavi et al. (1982); 3, Weydert et al. (1985); 4, Arnold et al. (1983); 5, Domdey et al. (1983); 6, Periasamy et al. (1984); 7, Nudel et al. (1984); 8, Zakut et al. (1982); 9, Weydert et al. (1983); 10, Robert et al. (1984); 11, Nabeshima et al. (1984); 12, Fornwald et al. (1982). Abbreviations: MLC, myosin light chain; MHC, myosin heavy chain; F, fast skeletal muscle; G3PDH, glyceraldehyde-3-phosphate dehydrogenase

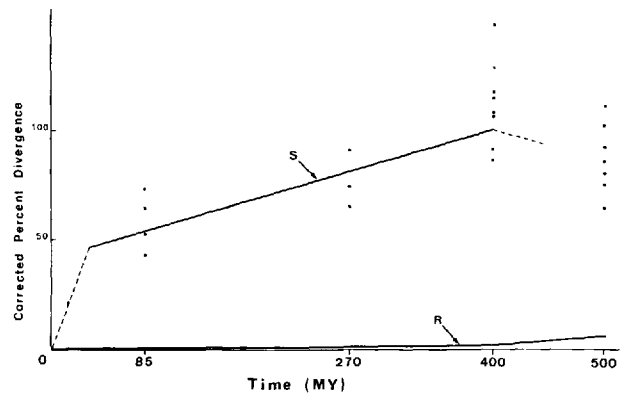


Fig. 3. Curves for divergence of actin-coding sequences. The corrected percentage divergences at silent (S) and replacement (R) sites (see Table 4) are plotted against divergence times in millions of years (MY), which are minimal estimates derived mainly from the fossil record (see text for discussion and references). Accumulation of silent substitutions begins to saturate between −30 and −40 MY

of their environments. There is a general tendency toward heavier isochore components in mammals, and particularly in humans. This is reflected in the biases in G + C content for actin-coding sequences in different species. The maintenance across species of differences in bias between actin genes would suggest that their distribution in heavier or lighter isochores tends to be conserved in evolution and occurred early during the evolution of warm-blooded vertebrates, or even prior to this, in cold-blooded vertebrates.

IV. Evolution of Actin-Coding Sequences

The accumulation of data on a large number of actin nucleotide sequences from widely different species (from yeast to human) now makes it possible to examine the substitution level between two actin sequences and to estimate the time since divergence of these sequences. We have used the method of calculating divergence corrected for multiple events between two homologous sequences that was introduced by Perler et al. (1980). When coding regions of two related genes are compared (codon by codon), two types of substitutions can be distinguished: those producing an amino acid change (replacement substitutions), and those leading to a synonymous codon (silent substitutions). In Perler's method of analysis, numbers of both replacement and silent substitutions are calculated, and are corrected for multiple substitutions at single sites using the "random substitution" model. The rate of divergence is expressed in unit evolutionary periods (UEP), which is the time in millions of years (MY) required for the fixation of a 1% substitution difference between two initially identical sequences. A distinction is made between UEP_R (the UEP for the replacement sites) and UEP_S (the UEP for the silent sites). The type of analysis introduced by Perler et al. (1980) and used here obviously has its limitations: For example, it does not take into account the possibility that some substitutions are more probable than others (see discussion in Cooke et al. 1981), nor the effect of sequence-homogenization events in paralogous comparisons.

Table 4 gives corrected percentage divergences for silent and replacement substitutions between different actin-coding sequences. Comparisons are classified into categories according to the estimated time at which either two species or two isoforms diverged. Paleontological data show the mammalian radiation to have occurred at approximately −85 MY (McKenna 1969, 1975; Romero-Herrera et al. 1973) and the divergence between mammals and birds to have occurred at approximately −270 MY (Dickerson 1971; Moore et al. 1976; Wilson et al. 1977). The gene duplication that led to the two sarcomeric actin genes (α_{sk} and α_c) probably oc-

curred during early amphibian evolution, 300–450 MY ago (Vandekerckhove and Weber 1984), and muscle-type actin isoforms exist only in vertebrates, which are estimated to have emerged 500-600 MY ago. On plots of corrected percentage divergence as defined by Perler et al. (1980) against divergence time, the UEP corresponds to the inverse of the slope (Fig. 3). The evolutionary-clock hypothesis predicts that the accumulation of replacement (or silent) substitutions is proportional to the divergence time. That an approximately linear plot is obtained as far back as 400 MY ago suggests that the initial assumptions, such as that of α_{sk}/α_c divergence during early amphibian evolution (Vandekerckhove and Weber 1984), are approximately correct. The partial sequence homogenization in the 5′ and 3′ regions is reflected in Table 4 by the lower percentages of silent substitutions in paralogous comparisons (i.e., mouse α_c/α_{sk}, human β/α_{sk}, and mouse β/α_{sk}) compared with orthologous comparisons. This phenomenon will inevitably reduce the precision with which the divergence time in paralogous comparisons can be estimated.

a. Replacement Sites. The extent of replacement substitution is about 1% between cardiac and skeletal muscle actin coding sequences and corresponds approximately to the degree of amino acid difference between the two actins (4/375). Evolution at replacement sites between the two sarcomeric actin genes is very slow, with a UEP_R of approximately 400 (i.e., 400 MY/1) since the divergence of the early amphibians. The extent of replacement substitution is about 5% for nonmuscle- and muscle-actin coding sequences, which is similar to the degree of amino acid difference (25/375).

Accumulation of replacement substitutions is estimated to have been more rapid in actin genes (by 8 to 16 fold) during early chordate and amphibian evolution than subsequently. Thus actin genes show an accelerated rate of divergence during evolution, and the fixation of substitutions at replacement sites does not behave as an accurate evolutionary clock in this case.

The rate of replacement substitutions is a function of the protein, and the UEP_R of actins (400) is close to those of tubulins (550) and histones H4 (400) or H3 (330) (Wilson et al. 1977). Other genes have evolved much more rapidly, like those encoding insulin chains A and B (14), globins (10), and the hypervariable region of fibrinopeptides (0.9) (Perler et al. 1980).

b. Silent Sites. Silent substitutions first accumulate rapidly in actin genes over a short initial period, the estimation of which in the case of the actins depends on the estimated divergence time between

Table 4. Corrected percentage divergences of actin-coding sequences

	Silent substitutions	Replacement substitutions
Rat/mouse separation (max. −15 MY)		
α_{sk} Rat/mouse	20	0.3
β Rat/mouse	23	0.3
Mammalian radiation (−85 MY)		
α_{sk} Rat/human	64	0.3
Mouse/human	73	0.3
α_c Mouse/human	53	0.3
β Rat/human	43	0.3
Mouse/human	52	0.3
Birds/Mammals (−270 MY)		
α_{sk} Chick/mouse	65	0.2
Chick/rat	75	0.2
Chick/man	91	0.2
Early amphibian evolution (approx. −400 MY)		
Human α_c/human α_{sk}	115	1
Human α_c/rat α_{sk}	108	1
Human α_c/mouse α_{sk}	92	1
Human α_c/chick α_{sk}	87	1
Mouse α_c/mouse α_{sk}	107	1
Mouse α_c/rat α_{sk}	129	1
Mouse α_c/chick α_{sk}	118	1
Mouse α_c/human α_{sk}	148	1
Early vertebrate evolution (between −500 and −600 MY)		
Human β/human α_{sk}	65	5
Human β/human α_c	93	5
Human β/mouse α_{sk}	81	5
Human β/mouse α_c	112	5
Human β/rat α_{sk}	76	5
Human β/chick α_{sk}	92	5
Rat β/rat α_{sk}	80	5
Rat β/mouse α_{sk}	76	5
Rat β/human α_{sk}	86	5
Rat β/chick α_{sk}	84	5
Rat β/mouse α_c	103	5
Rat β/human α_c	82	5
Mouse β/mouse α_{sk}	74	5
Mouse β/mouse α_c	104	5
Mouse β/rat α_{sk}	81	5
Mouse β/human α_{sk}	89	5
Mouse β/human α_c	79	5
Mouse β/chick α_{sk}	87	5

Corrected percentage divergences for silent- and replacement-substitution sites in each pair of actin-coding sequences were calculated as described by Perler et al. (1980). Comparisons are classified into categories according to the estimated time at which either two species or two isoforms diverged. In comparisons including mouse actin-coding sequences, we excluded from the calculations the first 39 codons. Sequence data for warm-blooded vertebrates were from the references in Table 2

rat and mouse. The corrected percentage divergence between these two species is nearly 20%, and it is unlikely that rat and mouse diverged more than 15 MY ago (Bonhomme F. personal communication). The initial slope in Fig. 3 is estimated based on this single point and corresponds to a maximum estimated UEP_S of 0.7 (i.e., 15 MY/20). Applying the same type of analysis to divergences between insulin genes and between globin genes, Perler et al. (1980) showed that changes at introns and at silent sites in coding regions first appear rapidly according to the same UEP_S value (0.7) as for the actins, which is therefore probably similar for different multigene families.

The initial accumulation of silent substitutions in actin genes begins to saturate between −30 and −40 MY (Fig. 3). Subsequently the rate of silent substitutions rapidly decreases, following a UEP_S of between 6 and 7, which again is in the same range as that for globins (10) and insulins (14). So the process of accumulation of silent substitutions may be similar for genes from different multigene families, whether they code for highly conserved proteins such as actins or for less conserved proteins such as globins.

If we extend the analysis beyond 400 MY ago, the accumulation of silent substitutions saturates, in contrast to that of replacement substitutions, leading to corrected percentage divergences between muscle- and nonmuscle-actin-coding sequences that are generally lower (65–112%) than those between cardiac and skeletal muscle actin coding sequences (87–148%) (see Table 4). This result was surprising, since the estimated divergence time between muscle- and nonmuscle-actin genes is greater (Vandekerckhove and Weber 1984) than that corresponding to the separation of the two striated actin genes (α_c and α_{sk}). In fact, however, this result is in keeping with the codon-usage tendency described above for paralogous comparisons (see Section III). Paralogous comparisons showed that restrictions on codon usage in β actin genes were intermediate between those in the two striated-actin genes, even though the latter sequences code for the most related proteins. This tendency seen within a single species appears to apply as well to comparisons between any warm-blooded vertebrate species. As already discussed, other phenomena, such as gene correction or the influence of isochore context (Bernardi et al. 1985) on codon-usage flexibility, intervene to distort the correlation between sequence divergence and evolutionary time.

Acknowledgments. We thank François Brégégère for the gift of pAL 41. Jean Michel Claverie and Bernard Caudron of the Institut Pasteur computer center kindly helped us with the computer programs. We also thank François Gros for support and encouragement.

This work was supported by grants to M.E. Buckingham from the French Ministry of Research and Technology (MRT), the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale, and the Muscular Dystrophy Association (MDA) of America. S. Alonso is grateful to the MRT and the Centre de Recherche Roussel-Uclaf for training fellowships.

References

Arnold HH, Krauskopf M, Siddiqui MAQ (1983) The nucleotide sequence of myosin light chain (L-2A) mRNA from embryonic chicken cardiac muscle tissue. Nucleic Acids Res 11: 1123–1131

Bagshaw CR (1982) Muscle contraction. In: Brammar WJ, Edidin M (eds) Outline studies in biology. Chapman and Hall, London New York, pp 30–38

Baltimore D (1981) Gene conversion: some implications for immunoglobulin genes. Cell 24:592–594

Bergsma DJ, Chang KS, Schwartz RJ (1985) Novel chicken actin gene: third cytoplasmic isoform. Mol Cell Biol 5:1151–1162

Bernardi G, Olofsson B, Filipski J, Zerial M, Salinas J, Cuny G, Meunier-Rotival M, Rodier F (1985) The mosaic genome of warm-blooded vertebrates. Science 228:953–958

Bernstein LB, Mount, SM, Weiner AM (1983) Pseudogenes for human small nuclear RNA U3 appear to arise by integration of self-primed reverse transcripts of the RNA into new chromosomal sites. Cell 32:461–472

Buckingham ME, Minty AJ (1983) Contractile protein genes. In: Maclean N, Gregory SP, Flavell RA (eds) Eukaryotic genes: their structure, activity and regulation, vol 21. Butterworths, London, pp 365–395

Challberg MD, Englund PT (1980) Specific labelling of 3′-termini with T_4 DNA polymerase. Methods Enzymol 65:39–43

Chang KS, Rothblum KN, Schwartz RJ (1985) The complete sequence of the chicken α-cardiac actin gene: a highly conserved vertebrate gene. Nucleic Acids Res 13:1223–1249

Claverie JM (1984) A common philosophy and FORTRAN 77 software package for implementing and searching sequence databases. Nucleic Acids Res 12:397–407

Cleveland DW, Lopata MA, MacDonald RJ, Cowan NJ, Rutter WJ, Kirschner MW (1980) Number and evolutionary conservation of α-and β-tubulin and cytoplasmic β- and γ-actin genes using specific cloned cDNA probes. Cell 20:95–105

Cooke NE, Coit D, Shine J, Baxter JD, Martial JA (1981) Human prolactin: cDNA structural analysis and evolutionary comparisons. J Biol Chem 256:4007–4016

Czosnek H, Nudel U, Mayer Y, Barker PE, Pravtcheva DD, Ruddle FH, Yaffé D (1983) The genes coding for the cardiac muscle actin, the skeletal muscle actin and the cytoplasmic β-actin are located on three different mouse chromosomes. EMBO J 2:1977–1979

Davidson EH, Britten RJ (1973) Organization, transcription and regulation in the animal genome. Rev Biol 48:565–613

Dickerson RE (1971) The structure of cytochrome c and the rates of molecular evolution. J Mol Evol 1:26–45

Domdey H, Wiebauer K, Klapthor H, Arnold HH (1983) Sequence analysis of the cloned mRNA coding for glyceraldehyde-3-phosphate dehydrogenase from chicken heart muscle. Eur J Biochem 131:129–135

Engel JN, Gunning PW, Kedes LM (1981) Isolation and characterization of human actin genes. Proc Natl Acad Sci USA 78:4674–4678

Fields S, Winter G (1981) Nucleotide sequence heterogeneity and sequence rearrangements in influenza virus cDNA. Gene 15:207–214

Fornwald JA, Kuncio G, Peng I, Ordahl CP (1982) The com-

plete nucleotide sequence of the chick α-actin gene and its evolutionary relationship to the actin gene family. Nucleic Acids Res 10:3861–3876

Fyrberg EA, Bond BJ, Hershey ND, Mixter KS, Davidson N (1981) The actin genes of *Drosophila*: Protein coding regions are highly conserved but intron positions are not. Cell 24: 107–116

Gallwitz D, Seidel R (1980) Molecular cloning of the actin gene from yeast *Saccharomyces cerevisiae*. Nucleic Acids Res 8: 1043–1059

Gunning P, Ponte P, Blau H, Kedes L (1983) α-Skeletal and α-cardiac actin genes are coexpressed in adult human skeletal muscle and heart. Mol Cell Biol 3:1985–1995

Gunning P, Ponte P, Kedes L, Hickey RJ, Skoultchi A (1984a) Expression of human cardiac actin in mouse L cells: A sarcomeric actin associates with a non muscle cytoskeleton. Cell 36:709–715

Gunning P, Ponte P, Kedes L, Eddy R, Shows T (1984b) Chromosomal location of the coexpressed human skeletal and cardiac actin genes. Proc Natl Acad Sci USA 81:1813–1817

Hamada H, Petrino MG, Kakunaga T (1982) Molecular structure and evolutionary origin of human cardiac actin gene. Proc Natl Acad Sci USA 79:5901–5905

Hanauer A, Levin M, Heilig R, Daegelen D, Kahn A, Mandel JL (1983) Isolation and characterization of cDNA clones for human skeletal muscle α actin. Nucleic Acids Res 11: 3503–3516

Hastings KEM, Emerson CP Jr (1983) Codon usage in muscle genes and liver genes. J Mol Evol 19:214–218

Humphries SE, Whittall R, Minty A, Buckingham ME, Williamson R (1981) There are approximately 20 actin genes in the human genome. Nucleic Acids Res 9:4895–4908

Kost TA, Theodorakis N, Hugues SH (1983) The nucleotide sequence of the chick cytoplasmic β-actin gene. Nucleic Acids Res 11:8287–8301

Kvist S, Brégégère F, Rask L, Cami B, Garoff H, Daniel F, Wiman K, Larhammar D, Abastado JP, Gachelin G, Peterson PA, Dobberstein B, Kourilsky P (1981) cDNA clone coding for part of a mouse H-2^d major histocompatibility antigen. Proc Natl Acad Sci USA 78:2772–2776

Mahdavi V, Periasamy M, Nadal-Ginard B (1982) Molecular characterization of two myosin heavy chain genes expressed in the adult heart. Nature 297:659–664

Maxam AM, Gilbert W (1980) Sequencing of end-labelled DNA with base specific cleavage reactions. Methods Enzymol 65: 499–560

Mayer Y, Czosnek H, Zeelon PE, Yaffé D, Nudel U (1984) Expression of the genes coding for the skeletal muscle and cardiac actins in the heart. Nucleic Acids Res 12:1087–1100

McKenna MG (1969) The origin and early differentiation of therian mammals. Ann NY Acad Sci 167:217–240

McKenna MG (1975) In: Luckett WP, Szalay FS (eds) Phylogeny of the primates. Plenum Press, New York, p 21

Minty AJ, Caravatti M, Robert B, Cohen A, Daubas P, Weydert A, Gros F, Buckingham ME (1981) Mouse actin mRNAs. Construction and characterization of a recombinant plasmid molecule containing a complementary DNA transcript of mouse α-actin mRNA. J Biol Chem 256:1008–1014

Minty AJ, Alonso S, Caravatti M, Buckingham ME (1982) A fetal skeletal muscle actin mRNA in the mouse and its identity with cardiac actin mRNA. Cell 30:185–192

Minty AJ, Alonso S, Guénet JL, Buckingham ME (1983) Number and organization of actin-related sequences in the mouse genome. J Mol Biol 167:77–101

Moore GW, Goodman M, Callahan C, Holmquist R, Moise H (1976) Stochastic versus augmented maximum parsimony method for estimating superimposed mutations in the divergent evolution of protein sequences. Methods tested on cytochrome c amino acid sequences. J Mol Biol 105:15–37

Moos M, Gallwitz D (1982) Structure of a human β-actin related pseudogene which lacks intervening sequences. Nucleic Acids Res 10:7843–7849

Nabeshima YI, Fujii-Kuriyama Y, Muramatsu M, Ogata K (1984) Alternative transcription and two modes of splicing result in two myosin light chains from one gene. Nature 308: 333–338

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

Nudel U, Zakut R, Shani M, Neuman S, Levy Z, Yaffé D (1983) The nucleotide sequence of the rat cytoplasmic β-actin gene. Nucleic Acids Res 11:1759–1771

Nudel U, Calvo JM, Shani M, Levy Z. (1984) The nucleotide sequence of a rat myosin light chain 2 gene. Nucleic Acids Res 12:7175–7186

Ordahl CP, Cooper TA (1983) Strong homology in promoter and 3′ untranslated regions of chick and rat α-actin genes. Nature 303:348–349

Periasamy M, Strehler EE, Garfinkel LI, Gubits RM, Ruiz-Opazo N, Nadal-Ginard B (1984) Fast skeletal muscle myosin light chains 1 and 3 are produced from a single gene by a combined process of differential RNA transcription and splicing. J Biol Chem 259:13595–13604

Perler F, Efstratiadis A, Lomedico P, Gilbert W, Kolodner R, Dodgson J (1980) The evolution of genes: the chicken preproinsulin gene. Cell 20:555–566

Pollard TD, Weihing RR (1974) Actin and myosin and cell movement. CRC Crit Rev Biochem 2:1–65

Ponte P, Gunning P, Blau H, Kedes L (1983) Human actin genes are single copy for α-skeletal and α-cardiac actin but multicopy for β-and γ-cytoskeletal genes: 3′ Untranslated regions are isotype specific but are conserved in evolution. Mol Cell Biol 3:1783–1791

Ponte P, Ng SY, Engel J, Gunning P, Kedes L (1984) Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human beta-actin cDNA. Nucleic Acids Res 12:1687–1696

Robert B, Weydert A, Caravatti M, Minty A, Cohen A, Daubas P, Gros F, Buckingham ME (1982) cDNA recombinant plasmid complementary to mRNAs for light chains 1 and 3 of mouse skeletal muscle myosin. Proc Natl Acad Sci USA 79: 2437–2441

Robert B, Daubas P, Akimenko MA, Cohen A, Garner I, Guénet JL, Buckingham ME (1984) A single locus in the mouse encodes both myosin light chains 1 and 3, a second locus corresponds to a related pseudogene. Cell 39:129–140

Robert B, Barton P, Minty A, Daubas P, Weydert A, Bonhomme F, Catalan J, Chazottes D, Guénet JL, Buckingham ME (1985) Investigation of genetic linkage between myosin and actin genes using an interspecific mouse back-cross. Nature 314: 181–183

Romero-Herrera AE, Lehman H, Joysey KA, Friday AE (1973) Molecular evolution of myoglobin and the fossil record: a phylogenetic synthesis. Nature 246:389–395

Sanger F, Coulson AR (1978) The use of thin acrylamide gels for DNA sequencing. FEBS Lett 87:107–110

Slightom JL, Blechl AE, Smithies O (1980) Human fetal ^{G}γ- and ^{A}γ-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell 21:627–638

Staden R, McLachlan AD (1982) Codon preference and its use in identifying protein coding regions in long DNA sequences. Nucleic Acids Res 10:141–156

Ueyama H, Hamada H, Battula N, Kakunaga T (1984) Structure of a human smooth muscle actin gene (aortic type) with a unique intron site. Mol Cell Biol 4:1073–1078

Vandekerckhove J, Weber K (1979) The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine

fast skeletal muscle and rabbit slow skeletal muscle. Differentiation 14:123–133

Vandekerckhove J, Weber K (1984) Chordate muscle actins differ distinctly from invertebrate muscle actins. The evolution of the different vertebrate muscle actins. J Mol Biol 179: 391–413

Vandekerckhove J, De Couet HG, Weber K (1983) In: dos Remedios CG, Barden JA (eds) Actin: structure and function in muscle and non-muscle cells. Academic, Sydney, Australia, pp. 241–248

Vandekerckhove J, Bugaisky G, Buckingham ME (1986) J Biol Chem, in press

Volckaert G, Tavernier J, Derynck R, Devos R, Fiers W (1981) Molecular mechanisms of nucleotide sequence rearrangements in cDNA clones of human fibroblast interferon mRNA. Gene 15:215–233

Wain-Hobson S, Nussinov R, Brown RJ, Sussman JL (1981) Preferential codon usage in genes. Gene 13:355–364

Weydert A, Daubas P, Caravatti M, Minty AJ, Bugaisky G, Cohen A, Robert B, Buckingham ME (1983) Identification of a recombinant plasmid coding for an adult fast myosin heavy chain from mouse skeletal muscle: detection of a sequential accumulation of mRNAs encoding different myosin heavy chain isoforms during skeletal muscle development in vivo. J Biol Chem 258:13867–13874

Weydert A, Daubas P, Lazaridis I, Barton P, Garner I, Leader DP, Bonhomme F, Catalan J, Simon D, Guénet JL, Gros F, Buckingham ME (1985) Genes for skeletal muscle myosin heavy chains are clustered and are not located on the same chromosome as a cardiac myosin heavy chain gene. Proc Natl Acad Sci USA 82: 7183–7187

Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. Annu Rev Biochem 46:573–639

Zakut R, Shani M, Givol D, Neuman S, Yaffé D, Nudel U (1982) The nucleotide sequence of the rat skeletal muscle actin gene. Nature 298:857–859

Zimmer EA, Martin SL, Beverly SM, Kan YW, Wilson AC (1980) Rapid duplication and loss of genes coding for the α chains of hemoglobin. Proc Natl Acad Sci USA 77:2158–2162

Received April 30, 1985/Revised October 28, 1985