MOLECULAR AND CELLULAR BIOLOGY, June 1982, p. 674–684
0270-7306/82/060674-11$02.00/0

Vol. 2, No. 6

Human Cytoplasmic Actin Proteins Are Encoded by a Multigene Family

JOANNE ENGEL, PETER GUNNING, AND LARRY KEDES*

Howard Hughes Medical Institute and Department of Medicine, Stanford University School of Medicine and Veterans Administration Medical Center, Palo Alto, California 94305

Received 6 January 1982/Accepted 26 February 1982

We characterized nine human actin genes that we isolated (Engel et al., Proc. Natl. Acad. Sci. U.S.A. 78:4674–4678, 1981) from a library of cloned human DNA. Measurements of the thermal stability of hybrids formed between each cloned actin gene and α-, β-, and γ-actin mRNA demonstrated that only one of the clones is most homologous to sarcomeric actin mRNA, whereas the remaining eight clones are most homologous to cytoplasmic actin mRNA. By the following criteria we show that these nine clones represent nine different actin gene loci rather than different alleles or different parts of a single gene: (i) the restriction enzyme maps of the coding regions are dissimilar; (ii) each clone contains sufficient coding region to encode all or most of an entire actin gene; and (iii) each clone contains sequences homologous to both the 5′ and 3′ ends of the coding region of a cloned chicken β-actin cDNA. We conclude, therefore, that the human cytoplasmic actin proteins are encoded by a multigene family.

The actins represent a group of proteins that is abundant in most eucaryotic cell types and that is highly conserved throughout its evolution. They participate in a wide variety of processes including muscle contraction, amoeboid movement, cytokinesis, and mitotic division (11, 21). Although a number of the lower eucaryotes appear to make use of only a single major actin protein, at least six distinguishable actin isoforms are utilized by mammals (31). The two mammalian cytoplasmic actins, β and γ, are found in most cells as components of the cytoskeleton (3). The amino acid sequences of these proteins are similar in all of the eucaryotes yet studied. The other four mammalian actins, which include the skeletal muscle α-actin, cardiac muscle actin, and two smooth muscle actins, are tissue specific and are more closely related to each other than to the cytoplasmic actins (31). This conservation of primary structure may reflect their common role in muscle contraction. Since the level of resolution in the amino acid sequencing studies can only detect isoforms present at greater than 5% of the total actin content, it is unknown whether minor actin isoforms exist as well.

Studies utilizing recombinant DNA technology have convincingly demonstrated that, with the exception of yeast (10, 18), actin is encoded by a multigene family (4, 6, 7, 9, 13, 14, 17, 24, 30). On the basis of protein analysis, this result is certainly to be expected in mammals. The surprising finding, however, has been the discovery in the slime mold of the expression of different actin genes encoding an identical major actin protein and of another encoding a minor variant actin not previously detected by protein analysis (16). Furthermore, Cleveland et al. (4) have shown that the chicken genome may encode several cytoplasmic γ-actin genes, although only one γ-actin protein has been identified in this organism. Thus the genomic organization and expression of actin genes are more complex than might be expected from the analysis of the actin proteins.

We recently isolated nine unique, nonallelic DNA clones from the human genome which contain coding regions for actin genes, and our analysis of genomic nitrocellulose blots suggests that there may be as many as 30 actin genes in the human genome (7; unpublished data). Since only six different actin proteins have been identified in mammals, it is of interest to know both the type of the actin encoded by and the coding capacity of these clones. To this end, we now report that these nine clones each include all or most of a different actin gene and that the human genome contains at least eight different actin genes coding for cytoplasmic actin proteins. (We use "gene" to indicate a region of DNA which is capable of encoding a protein without implying that this DNA sequence is functional or ever expressed within a cell.)

MATERIALS AND METHODS

RNA and DNA isolation. RNA was isolated from HeLa cells or from human muscle tissue by standard methods as reported previously (7). Plasmid DNA,

Charon 4 phage DNA, and human high-molecular-weight nuclear DNA were prepared by the method of Engel (7). Restriction fragments prepared by digestion of plasmids were purified by electrophoresis through agarose gels in 89 mM Tris–89 mM borate–2.5 mM EDTA and eluted by binding to powdered silica glass (32). Briefly, each gram of agarose gel containing a restriction fragment was homogenized in 3 ml of NaI solution (containing 90.8 g of NaL and 1.5 g of Na_2SO_3 in 100 ml of water) and allowed to dissolve at 37°C for 1 to 2 h. Twenty microliters of a 50% glass slurry solution was added to 15 μg of DNA and shaken vigorously at 4°C for 1 to 12 h. The solution was pelleted by centrifugation and washed twice with 100 volumes of NaI solution followed by two washes (100 volumes each) with a 50% ethanol–100 mM NaCl–1 mM EDTA–10 mM Tris-hydrochloride (pH 9.5) solution. The DNA was then eluted twice in 10 volumes of 50 mM Tris-hydrochloride (pH 9.0)–0.2 M NaCl and incubated at 60°C for 20 min each time. The eluted DNA was precipitated by adding sodium acetate to 0.3 M plus 2.5 volumes of ethanol.

Purification of human actin mRNA. The *Eco*RI fragment, which encodes actin from clone λHRL34, was linked to ABM-cellulose (Miles Biochemicals) by standard methods (19). Total HeLa RNA (6 mg) was hybridized to 100 μg of DNA-cellulose. The specifically bound RNA was eluted into water and passed over oligo(deoxyribosylthymine) cellulose (Collaborative Research, Inc.). The eluate containing polyadenylic acid-containing RNA was lyophilized, dissolved in 10 μl of water, and ethanol precipitated. This preparation migrated as a single band of 2,000 to 2,100 nucleotides on a 7 M urea–5% polyacrylamide gel as determined by ethidium bromide staining and autofluorography of the ^{3}H-labeled mRNA (data not shown). Analysis by electrophoresis on two-dimensional gels of the protein products, which were synthesized in vitro with this mRNA preparation as the template, revealed equal amounts of β- and γ-actin protein (data not shown).

Nick translation and end labeling of DNA and RNA. Nick translation was carried out by the method of Rigby et al. (23). Plasmid DNA or human high-molecular-weight nuclear DNA was labeled to a specific activity of 10^8 cpm/μg. Plasmid DNA or purified actin mRNA was end labeled with [γ-^{32}P]ATP (Amersham Corp., 4,500 Ci/mmol) by polynucleotide kinase (New England Nuclear Corp.) by the method of Maxam and Gilbert (15). To prevent labeling of RNA that contaminated the DNA preparations, the DNA was digested with the appropriate enzyme in the presence of bacterial alkaline phosphatase (Miles Biochemicals), electrophoresed on preparative agarose gels, eluted by binding to silica glass, radiolabeled with polynucleotide kinase, digested with a second restriction enzyme, and then electrophoresed on a second preparative agarose gel. The fragments of interest were eluted again by binding to silica glass.

Restriction enzyme digestion and mapping, agarose gel electrophoresis, and transfer of nucleic acids to nitrocellulose paper. Purified DNA was digested with restriction enzymes under conditions specified by the supplier (New England Biolabs) by using a two- to threefold excess of enzyme. DNA fragments were separated by electrophoresis on agarose gels in 89 mM Tris–89 mM borate–2.5 mM EDTA and transferred onto nitrocellulose filter paper (Millipore Corp.) by the method of Southern (28). After prehybridization for 6 h at 65°C in 4× SSC (1× SSC = 0.15 M NaCl plus 0.015 M sodium citrate)–2.5× Denhardt's solution–50 mM sodium orthophosphate (pH 7.0), the blots were hybridized to 10^5 to 10^6 cpm per ml of nick-translated DNA probe in the above buffer for 18 h. Extensive washing of the filters was carried out in 0.5× SSC–0.1% sodium dodecyl sulfate at 65°C. Restriction enzyme sites were mapped by single and double restriction enzyme digests or by partial digestion of 5′ end-labeled fragments (27).

Subcloning of genomic actin sequences into plasmid vectors. λHRL actin genes containing isolates were digested with *Eco*RI, ligated to *Eco*RI-cleaved pBR325 (1), and transfected into C600. Recombinants were screened for antibiotic resistance, and DNA was prepared from 5-ml overnight cultures of clones resistant to ampicillin and tetracycline and sensitive to chloramphenicol by a modification of the cleared lysate method (29). After digestion with *Eco*RI, the recombinant plasmid DNA was electrophoresed on 0.8% agarose gels in 89 mM Tris–89 mM borate–2.5 mM EDTA. The subcloned *Eco*RI fragment was identified by size or by transfer of the DNA to nitrocellulose paper followed by hybridization to a DNA probe derived by nick translation of the actin-encoding insert of the chicken β-actin cDNA clone, or by both.

RNA-DNA thermal stability experiments. Positive mRNA selection was performed essentially as described previously (7), except that the RNA was eluted in 0.1× SSC by incubating at 55°C and by increasing the temperature stepwise in 5°C increments, allowing 5 min for temperature equilibration at each temperature step. After ethanol precipitation in the presence of 5 μg of carrier tRNA, the entire sample from each temperature was translated in reticulocyte lysate mix (Amersham Corp.) and electrophoresed on one-dimensional and two-dimensional gels as reported previously (7).

RESULTS

Determination of the type of actin encoded by human actin gene clones. To identify the type of actin protein encoded by each of the nine human actin clones that we had previously isolated, we determined the thermal stability of hybrids formed between a particular clone and various human actin mRNAs. DNA isolated from each clone was bound to nitrocellulose filters and hybridized to a mixture which included α-, β-, and γ-actin mRNAs. Specifically bound RNA was then eluted at increasing temperatures and used as a template to direct the synthesis of proteins in a rabbit reticulocyte lysate. A sample of the [^{35}S]methionine-labeled protein products, synthesized in vitro with the mRNA template being eluted at each temperature step, was electrophoresed on a sodium dodecyl sulfate-polyacrylamide gel. Those samples containing labeled actin protein were further analyzed on a two-dimensional gel system capable of separating the α-, β-, and γ-actin proteins (20).

This protocol allowed us to compare the relative thermal stabilities of hybrids formed be-

tween each DNA clone and different actin mRNAs. Three types of melting behavior were observed among the nine clones examined, and an example of each type is shown in Fig. 1. Column A shows that all three actin mRNAs could be eluted from λHRL83 at low temperatures, but that only α-actin mRNA was present in the eluant at high temperatures. The data presented in Fig. 1B suggest that λHRL24 encodes a β-like actin; only β-actin mRNA was bound to it and was subsequently eluted at all temperatures tested. Fig. 1C shows results obtained with a class of clones that hybridized poorly to α-actin mRNA, but that released β- and γ-actin mRNA at a constant ratio at all temperatures tested.

An analysis of such experiments for each of the nine clones reveals the surprising result that eight of these clones are most homologous to the cytoplasmic β- and γ-actin mRNAs (Table 1). Only one of these clones (λHRL83) encoded an α-actin-like protein (Fig. 1, Table 1). It should be noted that we did not include actin mRNA from cardiac or from smooth muscle tissue in the hybridization mixture. Since the muscle-specific actin proteins differ from each other by only a few amino acids, one might expect that the mRNAs that encode them are also well conserved. However, Shani and co-workers (26) have found that rat smooth muscle actin mRNA has diverged significantly from both the rat skeletal muscle and cytoplasmic actin genes. Thus, one would expect a smooth muscle actin gene to hybridize similarly to α-, β-, and γ-actin mRNA. In contrast, the rat cardiac muscle mRNA is still quite homologous to the rat skeletal muscle actin gene as assayed by thermal stability or by S1 protection experiments of mRNA-DNA hybrids (26). Similar results have been reported in the chicken actin gene studies (25). If human smooth muscle actin genes and skeletal muscle actin genes have also diverged significantly from each other in DNA sequence, then we can expect that the hybridization of α-actin mRNA to λHRL83 at high temperatures would allow us to identify that clone as encoding a cardiac or an α-actin protein, rather than a smooth muscle actin isoform. Preliminary DNA sequencing in fact indicates that this clone encodes cardiac actin-specific amino acid residues at each of the locations known to distinguish it from other actin isoforms (amino acids 1 through 4, 298, and 357; H. Hamada and T. Kakunaga, personal communication). Four clones (λHRL21, λHRL24, λHRL25, and λHRL35) hybridize preferentially or solely to β-actin mRNA. We have confirmed the β-actin-like identity of clones λHRL21 and λHRL25 by DNA sequence analysis (unpublished data). The remaining four isolates, λHRL23, λHRL34, λHRL45, and λHRL84, hybridize to both β- and γ-actin mRNA at stringent temperatures. Although λHRL45 appears to release γ-actin mRNA only at the higher temperatures, the level of γ-actin mRNA in the mixtures of RNAs used in the hybridization was about fivefold less than that of β-actin mRNA, making detection of γ-actin mRNA difficult at low efficiencies of hybridization. λHRL34 hybridizes to almost equimolar amounts of β- and γ-actin mRNA, thus enriching for γ- over β-actin mRNA (Fig. 1 and data not shown). These two clones, therefore, may encode γ-actin proteins.

The thermal stability of the actin clone-mRNA hybrids at high temperatures implies that each of these cloned sequences is very homologous to its preferentially bound mRNA. It should be noted that the recombinants which hybridized to only β-actin mRNA have apparent melting temperatures approximately 5°C higher than those which hybridize to both β- and γ-actin mRNA. This result may indicate that considerable sequence differences in the untranslated regions or third base changes in the codons (or both) distinguish β- from γ-actin mRNAs. Similarly, the lack of α-actin mRNA hybridization to the cytoplasmic actin clones would suggest that α-actin mRNA is diverged significantly from β- and γ-actin mRNAs, a result predicted by amino acid comparisons of the proteins (3).

Since only two cytoplasmic actin proteins (β and γ) have been identified in mammals (3), the isolation of eight clones that encode cytoplasmic actin proteins is of considerable interest. There are three possible explanations for these data. First, although the recombinant DNA library from which these clones were selected was constructed from the DNA of a single individual, some of these clones could still be allelic and, therefore, not represent eight different actin coding loci. Second, some of these clones may encode different regions of the same gene (which would have to be spread over many kilobases [7]). Finally, and of most interest, they could represent eight different genes, each of which may be capable of encoding a cytoplasmic actin protein. To distinguish between these possibilities, we performed the experiments described below.

Mapping actin coding and flanking sequences. We previously used hybridization to cloned chicken actin cDNA and *Drosophila* genomic actin gene probes to demonstrate that each of these human actin gene clones encodes actin coding sequences on only a single *Eco*RI fragment (7). We have now corroborated these results by using a homologous actin gene probe prepared from human actin mRNA (see Materials and Methods). Purified human β- and γ-actin mRNA was fragmented, radiolabeled at the 5′

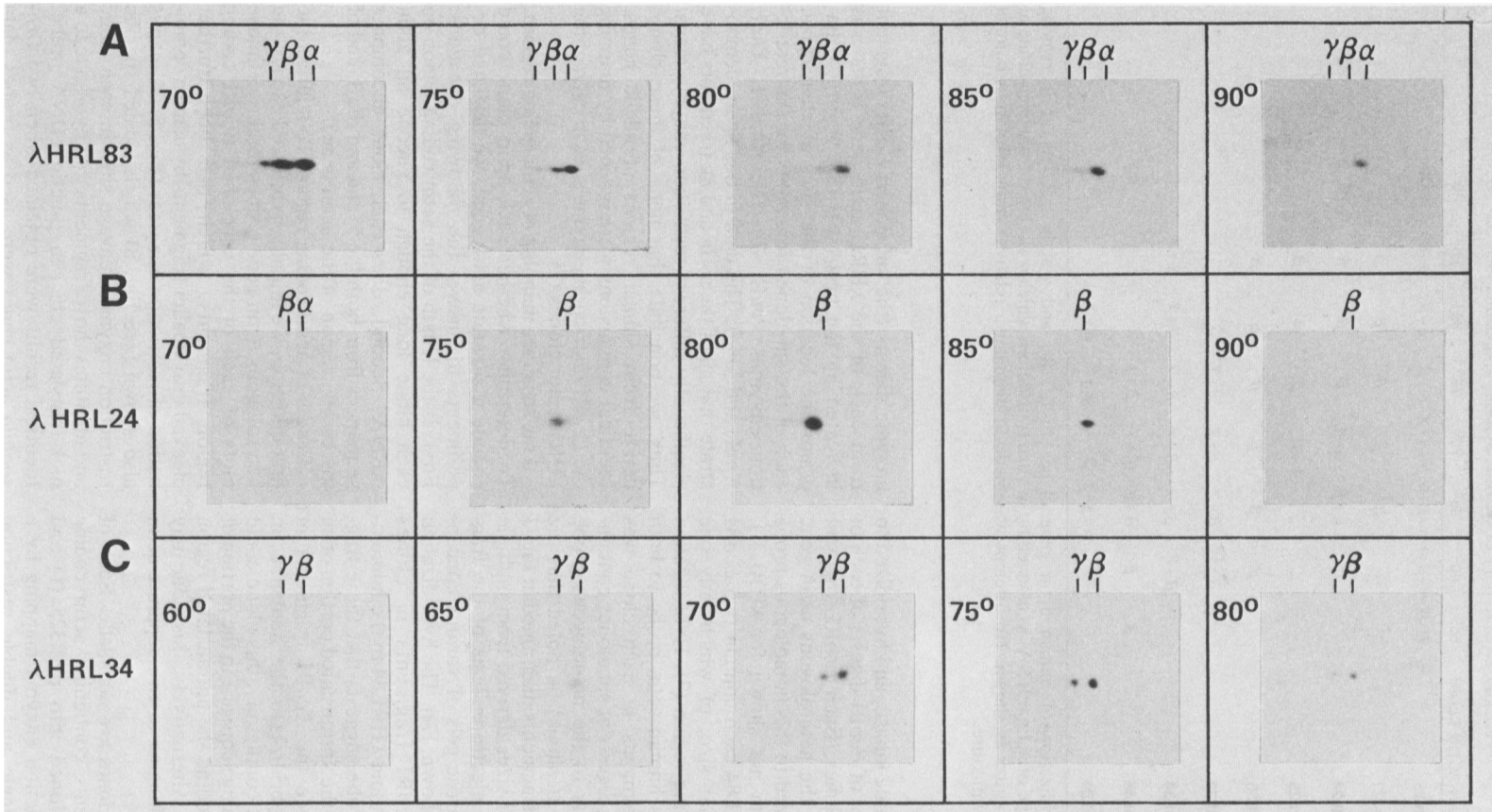

FIG. 1. Comparisons of the thermal stability of RNA-DNA hybrids formed between different actin mRNAs and cloned actin genes. DNA (20 μg) from human actin gene clone λHRL83 (A), λHRL24 (B), or λHRL34 (C) was bound to nitrocellulose filters and hybridized to total RNA isolated from HeLa cells and human muscle tissue. The specifically bound RNA was eluted by increasing the temperature in 5°C increments from 55°C to 90°C. RNA eluted at each temperature was translated in reticulocyte lysates, and the [^{35}S]methionine labeled protein products were separated by electrophoresis on two-dimensional gels. The actin-containing region of each gel was identified by Coomassie blue staining of rat myoblast proteins coelectrophoresed with each translation sample. An autofluorogram of this part of each gel is shown in each block. The elution temperature and the type of actin protein are indicated. Gels were exposed for 1 week to 2 months.

TABLE 1. Summary of thermal stability experiments[a]

HUMAN ACTIN CLONE	55°	60°	65°	70°	75°	80°	85°	90°
λHRL83	α+β+γ	α+β+γ	α+β+γ	α+β+γ	α+β	α+β	α+β	α
λHRL21			β	β	β	β	β	
λHRL24			α+β	α+β	β	β	β	β
λHRL25				β	β	β	β	
λHRL35		β	β	β	β	β	β	
λHRL23		β+γ	β+γ	β+γ	β+γ	β+γ		
λHRL34		β+γ	β+γ	β+γ	β+γ	β+γ	β+γ	
λHRL45	β	β	β+γ	β+γ	β+γ			
λHRL84			α+β+γ	β+γ	β+γ	β+γ		

[a] Differential hybridization of human actin gene clones to α-, β-, and γ-actin mRNAs at increasing temperatures. The type of actin mRNA, as assayed by in vitro translation, which was eluted from each clone at the indicated temperatures is shown. No attempt was made to distinguish relative amounts of translational activity of each assay mixture.

end by polynucleotide kinase, and hybridized to a nitrocellulose blot of *Eco*RI digests of each of the λHRL actin clones. Each *Eco*RI fragment that hybridized to the heterologous actin gene probe also hybridized to this homologous probe (Fig. 2 and 3; data not shown for λHRL21). λHRL25 and λHRL84 each contain two *Eco*RI fragments of similar size, of which only one contains actin coding region (7). Hybridization of the homologous human actin probe to both similarly sized fragments in each clone was verified by blotting digests of subclones containing these fragments to the homologous probe (data not shown). It should be noted that the sizes of the *Eco*RI fragments diagramed in Fig. 3 are the averages of sizes derived from multiple determinations by electrophoresis of the fragments through agarose gels of various concentrations (data not shown; [7]). The hybridization of additional *Eco*RI fragments in clones λHRL23, λHRL25, and λHRL84 to the homologous actin gene probe suggests that these fragments may specify the untranslated portion of an actin mRNA. Since the *Eco*RI fragments in λHRL23 which hybridize to the human actin gene probe are not contiguous, this clone could contain an intervening sequence in the portion of the actin gene encoding the untranslated region of its actin mRNA. Alternatively, this clone may encode regions belonging to two separate actin genes.

The actin gene clones are nonallelic. Each of the *Eco*RI fragments containing actin coding regions was subcloned into pBR325 (1) and subjected to restriction enzyme mapping by a combination of single and double restriction enzyme digests (the subcloned *Eco*RI fragment from each of the λHRL actin gene isolates is indicated by the prefix pHRL). Figure 4 diagrams the location of the sites of cleavage in each of the subcloned fragments for the restriction enzymes *Pvu*II, *Bam*HI, *Hin*dIII, *Xho*I, *Kpn*I, and *Sma*I. The data in this figure demonstrate that all of the actin coding fragments have significantly different restriction enzyme maps. Thus, we conclude that none of these clones derive from similar regions of allelic genes. Identical conclusions were reached by mapping the *Eco*RI restriction enzyme sites in each of the λHRL actin clones (7).

Fine structure mapping of actin coding regions. To determine whether each actin clone could encode a separate actin gene, we designed experiments to answer the following questions. First, is the length of the actin coding region on each clone long enough to encode an actin mRNA? Second, does each clone encompass sequences from both the 5′ end and the 3′ end of the coding region of the actin gene?

Length of actin coding region. To address the first question, we initially hybridized each of the heterologous actin gene probes to nitrocellulose blots of each of the subcloned *Eco*RI coding region fragments, which had been further cleaved into smaller fragments by using a combination of restriction enzymes (Fig. 5). A map is also shown (see Fig. 8) which locates specific restriction enzyme cleavage sites in each clone and identifies those fragments that hybridize to a nick-translated chicken β-actin cDNA probe. Identical results were obtained when two *Drosophila* actin genes were used as probes (data

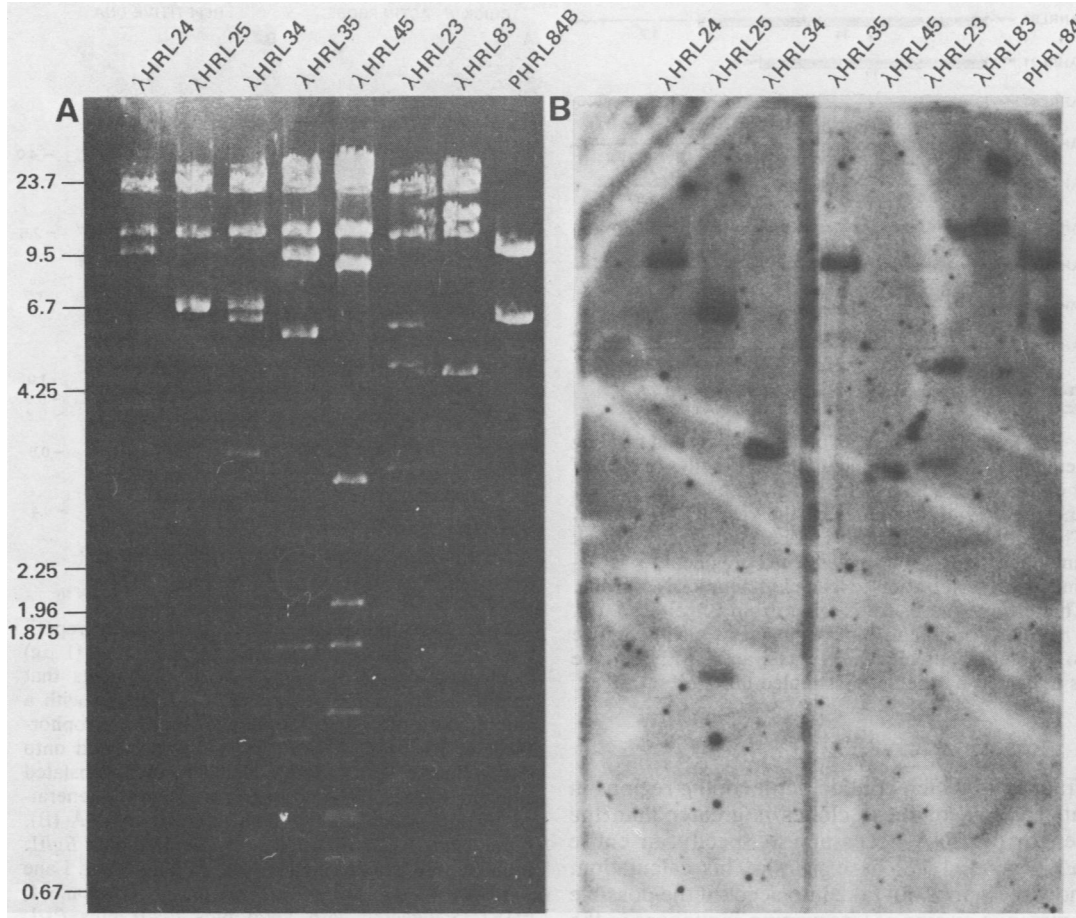

FIG. 2. Identification of actin coding and flanking sequences. DNA (1 μg) from each of the λHRL actin gene clones (λHRL24, λHRL25, λHRL34, λHRL35, λHRL45, λHRL23, or λHRL83) or subcloned *Eco*RI fragments (pHRL84B, a subclone which contains the 9.8-kilobase *Eco*RI actin coding fragment of λHRL84) was digested with *Eco*RI and electrophoresed through a 0.6% agarose gel. The ethidium bromide stain of the gel is shown in (A). (B) Gel was blotted onto nitrocellulose paper (28) and hybridized to 10^6 cpm of ^{32}P-labeled purified human β- and γ-actin mRNA. Sizes in kilobases are indicated on the left. The observed hybridization to the 4.6-kilobase *Eco*RI vector fragment of pHRL84B resulted from vector contamination of the probe during purification of the human β- and γ-actin mRNA (see the text).

not shown). The hybridization observed only corresponds to actin coding and perhaps 5′ untranslated sequences. The latter is unlikely, however (26), since the 3′ untranslated region of the chicken β-actin cDNA does not cross-hybridize with human actin genes (4).

The length of the coding region and the nonallelic nature of each of the actin gene isolates were further evaluated by digesting each of the λHRL actin clones with the more frequently cleaving restriction enzymes *Msp*I, *Alu*I, or *Hin*fI. Nitrocellulose blots of these digests hybridized to a chicken β-actin cDNA probe show that six of the clones (λHRL23, λHRL25, λHRL34, λHRL35, λHRL82, and λHRL83) contain enough coding region to specify an entire actin gene (Fig. 6 and data not shown). By these criteria λHRL21, λHRL45, and λHRL84 do not encode an entire gene. However, it is possible that we are underestimating the coding capacity of these three isolates because fragments that are less than 100 base pairs in length would not be transferred efficiently in our nitrocellulose blot procedure. Furthermore, two fragments of identical size would be scored only as a single fragment. Evidence presented below suggests that these three clones do each contain most of an actin coding region. We conclude, then, that to the resolution limit of our restriction enzyme maps, the sum of the lengths of the restriction

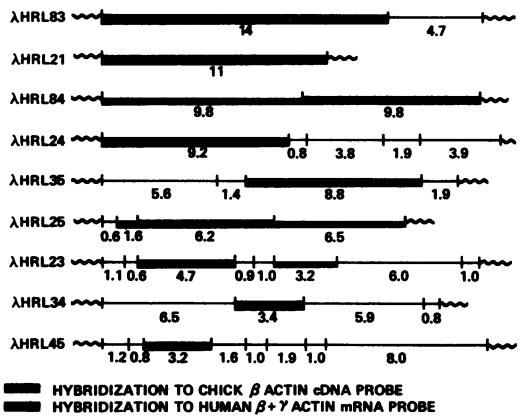

FIG. 3. Schematic map of actin coding and flanking sequences in λHRL clones. The *Eco*RI sites in each of the λHRL actin clones are indicated by vertical lines, and the wavy lines represent the vector CH4A arms. The actin coding region in each clone (heavy black line), localized to a single *Eco*RI fragment, was defined by hybridization to a chick β-actin cDNA probe. This part of the figure is taken from Engel et al. (7). The location of actin gene sequences which hybridize to a homologous human β- plus γ-actin mRNA probe is indicated by the heavy stippled line.

fragments which contain actin coding regions in at least six of these clones is greater than the length of DNA necessary to specify an entire actin gene (1,122 base pairs). This calculation, however, does not take into account the possible existence or size of intervening sequences in the coding regions of actin genes.

None of the putative cytoplasmic actin-encoding clones contained similarly sized *Msp*I, *Alu*I,

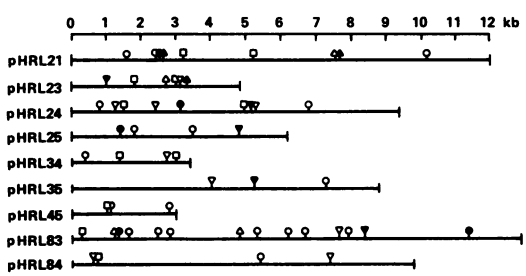

FIG. 4. Restriction enzyme maps of human actin gene clones. The location of the sites of cleavage of *Pvu*II (○), *Xho*I (●), *Hin*dIII (□), *Bam*HI (△), *Kpn*I (▲), *Sma*I (▼), and *Hpa*I (▽) are shown for each of the *Eco*RI actin coding regions containing fragments subcloned from each of the λHRL actin genes. An enzyme site that does not appear on a particular map indicates that this enzyme does not cut within this *Eco*RI fragment. Distances in kilobases are indicated by the scale at the top of the figure.

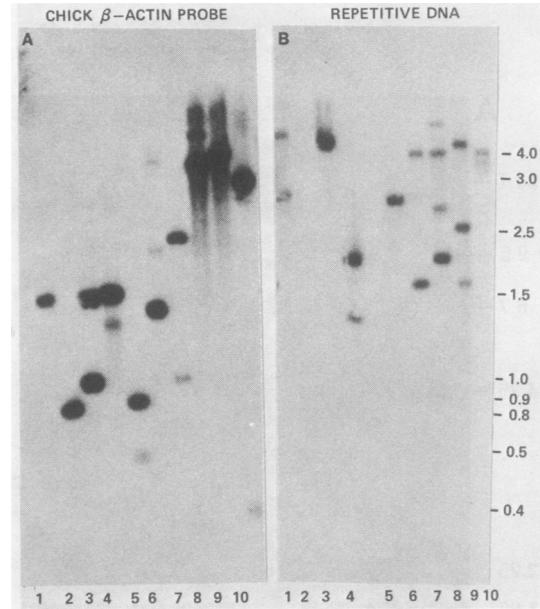

FIG. 5. Hybridization of coding region or repetitive DNA probe to human actin genes. DNA (1 μg) from each of the subcloned *Eco*RI fragments that contain actin coding sequences was digested with a combination of restriction enzymes and electrophoresed on 1% agarose gels. The gel was blotted onto nitrocellulose (28) and hybridized to nick-translated chicken β-actin cDNA probe (A) or to a probe generated by nick translation of human nuclear DNA (B). Lane 1, pHRL21 digested with *Hin*dIII plus *Bgl*II. Lane 2, pHRL45 digested with *Bgl*II plus *Pvu*II. Lane 3, pHRL 84 digested with *Hpa*I plus *Bgl*II. Lane 4, pHRL25 digested with *Sma*I plus *Pvu*II plus *Pst*I. Lane 5, pHRL34 digested with *Hin*cII plus *Hin*dIII. Lane 6, pHRL23 digested with *Kpn*I plus *Bam*HI. The hybridization of the chicken β-actin probe to the 2- and 4-kilobase fragments resulted from incomplete digestion of the DNA. Lane 7, pHRL35 digested with *Hpa*I plus *Pvu*II plus *Bgl*II. Lane 8, pHRL24 digested with *Sma*I plus *Pvu*II. Lane 9, pHRL24 digested with *Hpa*I. Lane 10, pHRL83 digested with *Bgl*II plus *Sma*I plus *Bam*HI. Sizes in kilobases are shown on the right.

or *Hin*fI fragments which hybridized to this coding region probe. From the DNA sequence heterogeneity of each of the coding regions of these actin genes suggested by this result, we can again infer the nonallelic nature of each of the eight cytoplasmic actin gene clones. The diversity of *Hin*fI sites contrasts their conservation in the actin genes of *Drosophila* (8) and sea urchins (24) and may reflect the presence of intervening sequences. We have shown previously (7) that this sequence divergence between different actin genes reflected in restriction enzyme site variation does not result from genetic polymorphism in the actin gene loci. Also, as

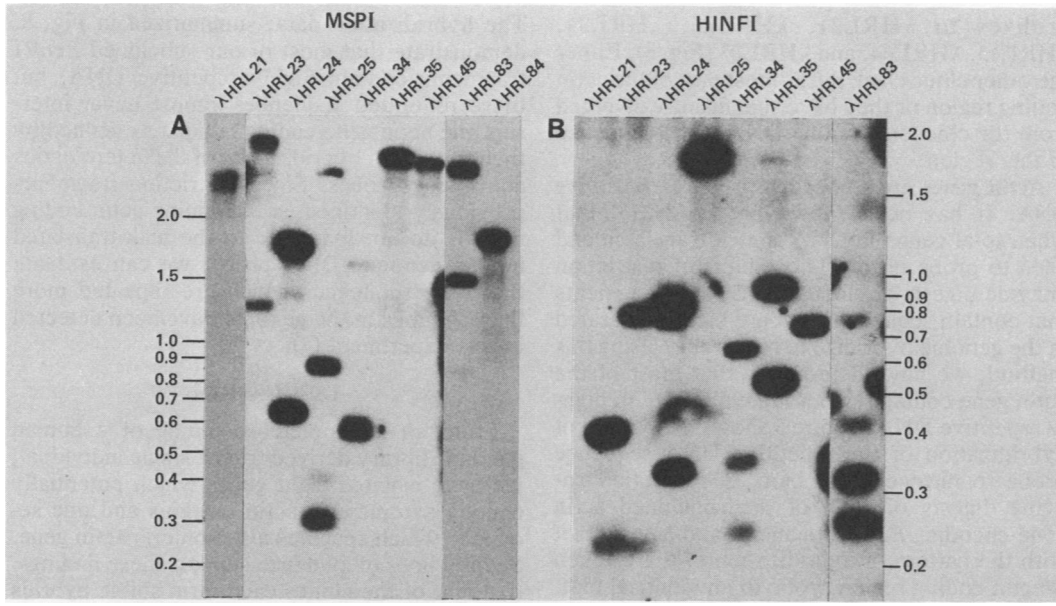

FIG. 6. Estimation of length of coding region contained in each human actin gene clone. DNA (2 μg) from each of the λHRL actin clones was digested with *Msp*I (A) or *Hin*fI (B) and electrophoresed on 2% agarose gels. The DNA was transferred to nitrocellulose (28) and hybridized to a nick-translated chicken β-actin cDNA probe. Sizes in kilobases are indicated to the left or right of the appropriate blot.

noted earlier, each of these actin genes that we have cloned derives from the DNA of a single individual.

Gene orientation. To further ascertain whether each of our actin clones could encode an entire gene, we made use of 5′-end and 3′-end coding region probes derived from the chicken β-actin cDNA clone. This heterologous probe was digested with *Bgl*I plus *Pst*I to yield a 0.5-kilobase fragment containing 300 bases from the 5′ untranslated region of the mRNA and the first 200 bases of the coding region. A 3′-end coding region probe was isolated by digesting the chicken β-actin cDNA clone with *Kpn*I to generate a fragment which included 300 bases of the coding region plus 200 bases from the 3′ untranslated region (unpublished data). We reasoned that if both of these 5′- and 3′-end coding region probes hybridized to a particular human actin gene clone, then that clone must contain most or all of an actin gene. Furthermore, all such clones hybridizing to both 5′ and 3′ probes must be distinct from one other; they could not be derived from different parts of the same gene because they would contain overlapping coding regions. It should be noted that this approach assumes that the entire chicken β-actin cDNA coding sequence is homologous to the human actin gene coding sequences and that its specificity derives from the short length of the coding region contained in the 5′ and 3′ probes.

The appropriate *Bgl*I-*Pst*I and *Kpn*I fragments of the chicken β-actin cDNA clones were eluted from a gel and nick translated. No cross-hybridization of the fragments to each other or to any of the DNA fragments located between the two end fragments was observed (data not shown). These radioactive probes were then hybridized to nitrocellulose blots of appropriate restriction enzyme digests of each of the subcloned *Eco*RI coding fragments. Each of the nine human actin gene clones that we have described hybridizes to both the 5′ and 3′ chicken actin gene probes (Fig. 7). Figure 8 summarizes the hybridization data from each of the clones. The data presented in these figures demonstrate that all nine clones contain part or all of a distinct actin gene. In addition, the data define the polarity of transcription relative to the restriction enzyme map for most of the genes. The fact that λHRL21, λHRL45, and λHRL84 hybridize to both the 5′ and 3′ probes suggests that our previous experiments (Fig. 5) underestimated the size of their actin coding regions. We can conclude, therefore, that the human genome encodes at least eight actin genes which could specify a cytoplasmic actin protein. In addition, it encodes at least one sarcomeric actin sequence.

We have also prepared a *Bal*II-*Pst*I fragment which contains only the 50 most 5′ bases from the coding region of the chicken β-actin cDNA clone (unpublished data). This probe only hy-

bridizes to λHRL21, λHRL25, λHRL34, λHRL45, λHRL64, and λHRL83 (Fig. 8). Either the other clones lack the 5'-most part of the actin coding region or they have significantly diverged from the chicken β-actin cDNA clone sequence in this region.

Actin genes are embedded in highly repetitive DNA. It has been shown previously (5) that when total genomic DNA is nick translated and used to probe nitrocellulose blots of restriction enzyme digests of cloned DNA, only fragments that contain sequences that are highly reiterated in the genome hybridize to this probe. Using this method, we have determined that most of the actin gene coding regions are embedded in highly repetitive DNA. Figure 5 shows the pattern of hybridization of the repetitive DNA sequence probe to nitrocellulose blots of restriction enzyme digests of each of the subcloned actin gene-encoding *Eco*RI fragments and compares it with the pattern of hybridization of the chicken β-actin coding region probe to an identical blot.

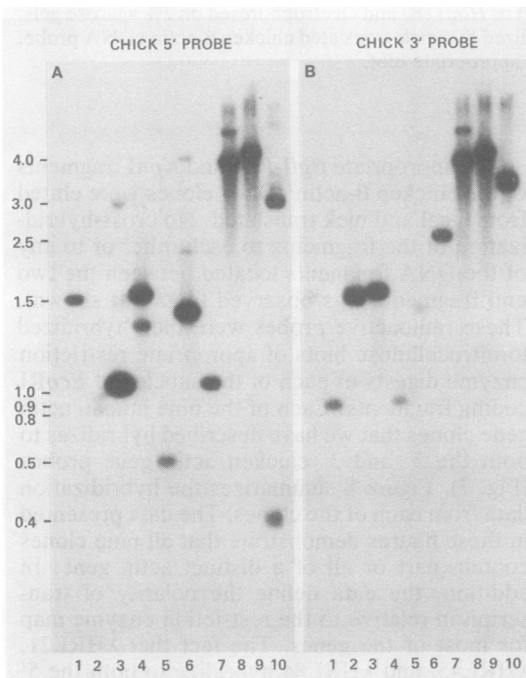

FIG. 7. Hybridization of 5' or 3' coding region probes to human actin genes. DNA (1 μg) from each of the subcloned *Eco*RI actin coding regions containing fragments was digested with a combination of restriction enzymes and electrophoresed on 1% agarose gels. The gel was blotted onto nitrocellulose (28) and hybridized to a nick-translated 5' coding region probe (*Pst*I-*Bgl*I fragment; A) or to a nick-translated 3' coding region probe (*Kpn*I fragment; B) derived from the chicken β-actin cDNA clone. The DNA and enzyme digests are the same as for Fig. 5. Sizes in kilobases are shown on the left.

The hybridization data, summarized in Fig. 8, demonstrate that most of our subcloned *Eco*RI fragments contain highly repetitive DNA, but these reiterated sequences almost never interrupt the actin gene coding regions as defined by their pattern of hybridization of the heterologous actin gene probes. Since restriction fragments previously identified as containing actin coding regions do not hybridize to the nick-translated human genomic DNA probe, we can estimate that only sequences which are repeated more than 30 times in the genome have been detected in this experiment (7).

DISCUSSION

From an incomplete screening of a human genomic library derived from a single individual, we have isolated eight genes which potentially encode cytoplasmic actin proteins and one sequence which specifies a sarcomeric actin gene. Several lines of evidence support these findings. (i) Eight of the clones each form stable hybrids with cytoplasmic actin mRNA but not with α-actin mRNA, whereas the ninth clone hybridizes most stably to α-actin mRNA. (ii) Restriction maps of the nine clones suggest that they are nonallelic. (iii) Each clone contains sufficient coding region, at the level of resolution afforded by the restriction maps, to encode all or most of a complete actin gene. (iv) Each clone contains sequences complementary to the first 200 nucleotides and to the last 300 nucleotides of the coding region of the chicken β-actin cDNA gene, thus indicating that none of them derive from different regions of the same gene.

These findings further suggest that the large number of actin genes found in the human genome (over 25 [7]) primarily encode cytoplasmic rather than muscle-specific actins. This conclusion must await further confirmation, however, since the human genomic library that we have screened does not necessarily reflect the relative sequence composition of the human genome with respect to actin genes (unpublished data). Because the library was constructed by cloning 15- to 20-kilobase fragments generated by a single partial *Eco*RI digestion, regions of the genome which have either very many or very few *Eco*RI sites are inevitably underrepresented. Thus, it remains a formal possibility that the muscle-specific actin genes have been selected against during the construction of this library.

Hamada and co-workers (12) have recently reported that only two bands of human nuclear DNA cleaved with *Eco*RI or *Bam*HI were observed to hybridize to probes generated by nick translation of a *Dictyostelium* actin cDNA or a human β-actin cDNA clone. Their results are at variance with our previously reported observation (7) of multiple *Eco*RI, *Bam*HI, or *Hin*dIII

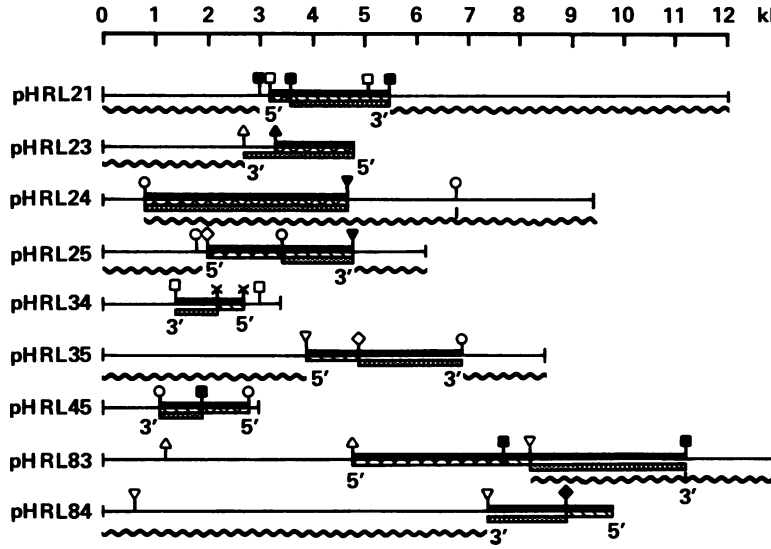

FIG. 8. Summary map of hybridization and restriction enzyme sites of cloned human actin genes. The hybridization of the chicken β-actin cDNA probe (▬), chicken β-actin *PstI-BglII* 5' coding region probe (▨), chicken β-actin *PstI-BglI* 5' coding region probe (▨), chicken β-actin *KpnI* 3' coding region probe (▨), or repetitive DNA probe (〰) is localized to fragments generated by restriction enzyme digestion of each subcloned *EcoRI* fragment (*PvuII* [○], *XhoI* [●], *BamHI*, [△], *KpnI* [▲], *HpaI* [▽], *SmaI* [▼], *HindIII* [□], *PstI* [◇], *BglII* [■], and *HincII* [✳]). The direction of transcription relative to the restriction map is indicated by marking the 5'-3' orientation of each clone. Distances in kilobases are shown by the scale at the top of the figure. The assignment of highly repetitive DNA sequences to the coding fragment in clone pHRL83 was confirmed by hybridizing subcloned restriction fragments of pHRL83 to the repetitive DNA sequence probe (data not shown).

fragments of human DNA that cross-hybridize to chicken or *Drosophila* actin gene probes and may reflect differences in technique.

It is now clear that the human cytoplasmic actins are encoded by a multigene family. Whether the genes that we have isolated are functional is unknown. Three of these isolates hybridize to a probe derived from the 3' untranslated region of a human γ-actin cDNA clone (unpublished data), a finding that suggests that they may in fact be transcribed. In the case of other multigene families such as the globin genes, the majority of these multiple gene copies are at least capable of being expressed (22). Recently, another group (2) has demonstrated the stable existence of a third isoform of cytoplasmic actin in a transformed mouse sarcoma cell line. This finding lends additional credence to the existence of multiple actin genes that code for cytoplasmic actin proteins in mammalian genomes.

Furthermore, the multiplicity of cytoplasmic actin genes is not unique to the human genome. DNA sequencing of portions from the six *Drosophila* actin genes has shown that they all more closely resemble cytoplasmic actin genes than skeletal muscle actin genes (8). In particular, one of these cytoplasmic-like actin genes encodes the major actin component of the insect flight muscle, thus demonstrating that the emergence of an α-actin-type protein occurred after the divergence of protostomes from deuterostomes. Similarly, *Dictyostelium* encodes 17 cytoplasmic actin genes (14). At least six of these genes are transcribed into mRNAs specifying different actin proteins. Three of these genes code for an identical protein, but are transcribed at different times during development (16). These data indicate that tissue and developmental stage-specific expression of cytoplasmic-like actin genes arose during evolution long before the advent of mammalian muscle-specific actin genes. Our finding of multiple human cytoplasmic actin genes raises the possibility that similar tissue-specific or developmental stage-specific expression of cytoplasmic actin genes may also occur in humans.

ACKNOWLEDGMENTS

We acknowledge the advice and encouragement of Rob Maxson, Tim Mohun, Glen Gormezano, Phyllis Ponte, and Alex Mauron. We thank Don Cleveland, Sally Tobin, and Eric Fyrberg for kindly donating actin gene probes.

L.K. is an Investigator and P.G. is a Research Associate of the Howard Hughes Medical Institute. J.E. was supported by the Medical Scientist Training Program.

LITERATURE CITED

1. **Bolivar, F.** 1978. Construction and characterization of new cloning vehicles: derivatives of plasmid pBR322

carrying unique *Eco*RI sites for selection of *Eco*RI-generated recombinant DNA molecules. Gene 4:121–136.
2. **Bravo, R., S. J. Fey, J. V. Small, P. M. Larsen, and J. E. Celis.** 1981. Coexistence of three major isoactins in a single sarcoma 180 cell. Cell 25:195–202.
3. **Clarke, M., and J. Spudich.** 1977. Non-muscle contractile proteins: the role of actin and myosin in cell motility and shape determination. Annu. Rev. Biochem. 46:797–822.
4. **Cleveland, D. W., M. A. Lopata, R. J. MacDonald, N. J. Cowan, W. J. Rutter, and M. W. Kirschner.** 1980. Number and evolutionary conservation of α- and β-tubulin and cytoplasmic β- and γ-actin genes using specific cDNA probes. Cell 20:95–105.
5. **Cochet, M., F. Gannon, R. Hen, L. Maroteaux, F. Perrin, and P. Chambon.** 1979. Organization and sequence studies of the 17-piece chicken conalbumin gene. Nature (London) 282:567–574.
6. **Durica, D. A., J. A. Schloss, and W. R. Crain.** 1980. Studies on the organization of actin gene sequences in the sea urchin: molecular cloning of an intron-containing sequence coding for a cytoplasmic actin. Proc. Natl. Acad. Sci. U.S.A. 77:5683–5687.
7. **Engel, J. N., P. W. Gunning, and L. H. Kedes.** 1981. Isolation and characterization of human actin genes. Proc. Natl. Acad. Sci. U.S.A. 78:4674–4678.
8. **Fyrberg, E. A., B. J. Bond, N. D. Hershey, K. S. Mixter, and N. Davidson.** 1981. The actin genes of drosophila: protein coding sequences are highly conserved but intron positions are not. Cell 24:107–116.
9. **Fyrberg, E. A., K. L. Kindle, N. Davidson, and A. Sodja.** 1980. The actin genes of Drosophila, a dispersed multigene family. Cell 19:365–378.
10. **Gallwitz, D., and I. Sures.** 1980. Structure of a split yeast gene: complete nucleotide sequence of the actin gene in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A. 77:2546–2550.
11. **Goldman, R., T. Pollard, and X. Rosenbaum (ed.).** 1976. Cell motility. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Hamada, H., J. Leavitt, and T. Kakunaga.** 1981. Mutated β-actin gene; coexpression with an unmutated allele in a chemically transformed human fibroblast cell line. Proc. Natl. Acad. Sci. U.S.A. 78:3634–3638.
13. **Humphries, S. E., R. Whittall, A. Minty, M. Buckingham, and R. Williamson.** 1981. There are approximately 20 actin genes in the human genome. Nucleic Acid Res. 9:4895–4908.
14. **Kindle, K. L., and R. A. Firtel.** 1978. Identification and analysis of dictyostelium actin genes, a family of moderately repeated genes. Cell 15:763–778.
15. **Maxam, A., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. 74:560–564.
16. **McKeown, M., and R. A. Firtel.** 1981. Differential expression and 5′ end mapping of actin genes in dictyostelium. Cell 24:799–807.
17. **McKeown, M., W. Taylor, K. Kindle, R. Firtel, W. Bender, and N. Davidson.** 1978. Multiple heterogeneous actin genes in dictyostelium. Cell 15:789–800.
18. **Ng, R., and J. Abelson.** 1977. Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A. 77:3912–3916.
19. **Noyes, B. E., and G. R. Stark.** 1975. Nucleic acid hybridization using DNA covalently coupled to cellulose. Cell 5:301–310.
20. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol.Chem. 250:4007–4021.
21. **Pollard, T. D., and R. R.Weihing.** 1974. Actin and myosin and cell movement. Crit. Rev. Biochem. 2:1–65.
22. **Proudfoot, N.** 1980. Pseudogenes. Nature (London) 286:840–841.
23. **Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling DNA to high specific activity *in vitro* by nick-translation with DNA polymerase I. J. Mol. Biol. 113:237–251.
24. **Scheller, R. H., L. B. McAllister, W. R. Crain, D. S. Durica, J. Posakony, T. L. Thomas, R. J. Britten, and E. H. Davidson.** 1981. Organization and expression of multiple actin genes in the sea urchin. Mol. Cell Biol. 1:609–628.
25. **Schwartz, R. J., and K. N. Rothblum.** 1981. Gene switching in myogenesis; differential expression of the actin multigene family. Biochem. 20:4122–4129.
26. **Shani, M., U. Nudel, D. Zevin-Sonkin, R. Zakut, D. Givol, D. Katcoff, Y. Carmon, J. Reiter, A. M. Frischaut, and D. Yaffe.** 1981. Skeletal muscle actin mRNA. Characterization of the 3′ untranslated region. Nucleic Acid Res. 9:579–589.
27. **Smith, H. O., and M. L. Birnstiel.** 1976. A simple method for DNA restriction site mapping. Nucleic Acid Res. 3:2387–2397.
28. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503–517.
29. **Sures, I., A. Maxam, R. H. Cohn, and L. H. Kedes.** 1976. Identification and location of the histone H2a and H3 genes by sequence analysis of sea urchin S. purpuratus DNA cloned in *E. coli*. Cell 9:495–502.
30. **Tobin S. L., E. Zulauf, F. Sanchez, E. A. Craig, and B. J. McCarthy.** 1980. Multiple actin-related sequences in the Drosophila melanogaster genome. Cell 19:121–131.
31. **Vandekerckhove, J., and K. Weber.** 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.
32. **Vogelstein, B., and D. Gillespie.** 1979. Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. U.S.A. 76:615–619.