

Isolation and characterization of human actin genes

(recombinant DNA/muscle proteins/multigene family/gene polymorphism)

JOANNE N. ENGEL, PETER W. GUNNING, AND LARRY KEDES*

Howard Hughes Medical Institute Laboratory and Department of Medicine at Stanford Medical School and Veterans Administration Medical Center, Palo Alto, California 94304

Communicated by L. L. Cavalli-Sforza, March 9, 1981

ABSTRACT We have utilized cloned actin genes from *Drosophila melanogaster* and from chicken to isolate 12 actin gene fragments from a human DNA library. Each of these 12 clones was shown to contain actin coding regions by its ability to selectively hybridize to human actin mRNA as assayed by *in vitro* translation. The translation product was judged to be actin on the basis of its comigration with authentic actins when electrophoresed on one- and two-dimensional NaDodSO₄/polyacrylamide gels and on the basis of its partial proteolysis products. Determination of the sizes and order of the fragments generated by restriction endonuclease digestion of each of these recombinant phages allows us to conclude that they are nonallelic and are from nonoverlapping regions of the genome. We have used these cloned human actin genes and the *Drosophila* and chicken actin gene clones to show that the human genome contains 25-30 *EcoRI* fragments homologous to actin genes and that, among three nonconsanguineous individuals tested, none of these fragments exhibit length or restriction-site polymorphism.

Actin is a highly conserved protein that is found in abundance in all eukaryotic cells and is functionally involved in cell motility, mitosis, muscle contraction, and the maintenance of cytoskeletal structure (1, 2). This diversity of roles within a cell is reflected by variant actins that differ in their isoelectric points (3). Amino acid sequence analysis of the mammalian actin proteins suggests that they are the products of at least six separate genes (4). The existence of multiple actin genes in several organisms has been confirmed most recently by direct isolation of the genes. Seventeen actin gene sequences, coding for at least two different actin proteins, have been identified in *Dictyostelium* (5, 6). The *Drosophila melanogaster* genome, on the other hand, contains only 6 actin genes specifying two or more actin species (7, 8), whereas the sea urchin genome codes for 5-20 actin genes (9).

A study of the structure and arrangement of the human actin genes in the genome is a prerequisite for understanding the developmental regulation of this multigene family in normal and diseased cells. In this paper we report the isolation and initial characterization of 12 different human genomic clones containing actin coding sequences. We use these cloned sequences and cloned actin genes from other species to determine the number of actin gene-containing fragments in the human genome and to search for the presence of actin gene polymorphism in the human population.

MATERIALS AND METHODS

RNA and DNA Isolation. RNA was prepared from human muscle by using the phenol/chloroform procedure of Palmiter (10). DNA and small molecular weight RNA was removed by

precipitation with 3 M sodium acetate (pH 6.0). RNA was isolated from HeLa cells grown in spinner flasks to a density of 10⁶ cells per ml by the guanidine hydrochloride method (11), followed by sodium acetate precipitation (10). Bacterial plasmid DNA was prepared as described (12). Charon phage recombinants were grown as plate stocks and recovered by CsCl gradient centrifugation (13). Human high molecular weight nuclear DNA was prepared from peripheral blood lymphocytes by the method of Kan *et al.* (14).

Nick-Translation. Nick-translation of the chicken β -actin cDNA clone (2000 base pairs long) (15), of the *D. melanogaster* actin clone coding for a cytoplasmic actin (which contains the entire gene except for the 5'-end 70 base pairs) (7), and of the *D. melanogaster* actin clone K1A, which codes for a muscle-specific actin (containing the entire coding region, including the intervening sequence) (8), (S. Tobin, personal communication) were carried out as described (16). The DNA was labeled to a specific activity of 10⁸ cpm per μ g.

Screening the Human Library. A bacteriophage λ library constructed from the partial *EcoRI*-digested DNA of a single human fetus was screened as detailed by Maniatis *et al.* (17) by using the *in situ* plaque hybridization technique (18). Plaques (400,000) were screened with the nick-translated chicken and *Drosophila* actin gene probes. After two rounds of plaque purification, 14 clones continued to hybridize to the three heterologous actin probes, and these isolates were chosen for further study.

Restriction Endonuclease Digestions, Agarose Gel Electrophoresis, and Transfer to Nitrocellulose Paper. Purified DNA was digested with restriction enzymes under conditions specified by the supplier (New England BioLabs) with a 2- to 3-fold excess of enzyme. DNA fragments were separated by electrophoresis on agarose gels in 89 mM Tris/89 mM borate/2.5 mM EDTA and were transferred onto nitrocellulose paper by the method of Southern (19). After prehybridization for 6 hr at 65°C in 0.60 M NaCl/0.060 M sodium citrate, pH 7.0/0.05% polyvinyl pyrrolidone/0.05% Ficoll/0.05% bovine serum albumin/50 mM NaPi, pH 7.0, the blots were hybridized to 10⁵ cpm per ml of nick-translated DNA probe in the above buffer for 18 hr. Extensive washing of the filters was carried out in 0.075 M NaCl/0.0075 M sodium citrate, pH 7.0/0.1% NaDodSO₄ at 65°C.

Positive mRNA Selection, *in Vitro* Translation, and NaDodSO₄/Polyacrylamide Gel Electrophoresis. Positive mRNA selection was carried out essentially as described (15). Cell-free translation of mRNAs that specifically hybridized to the cloned DNA fragments was performed with rabbit reticulocyte lysate (20) mix (Amersham) containing 1 mCi (1 Ci = 3.7 \times 10¹⁰ becquerels) per/ml of [³⁵S]methionine (New England Nuclear) per ml. Translation products were analyzed on one-dimensional polyacrylamide gels (21) and two-dimensional iso-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

* To whom reprint requests should be addressed.

electric focusing NaDodSO₄/polyacrylamide gels (22) and were detected by fluorography (23).

Analysis of Proteins by Partial Proteolysis. The radiolabeled band that was synthesized in the reticulocyte lysate mix programmed with actin mRNA purified by positive mRNA selection and that comigrated with rabbit α -actin was cut out of a NaDodSO₄/polyacrylamide gel and partially digested with *Staphylococcus aureus* V8 protease (Sigma) as described (24). The partial proteolysis products were electrophoresed on NaDodSO₄/15% (wt/vol) polyacrylamide gels.

Biosafety. This research was carried out in accordance with the National Institutes of Health guidelines with P2/EK2 containment.

RESULTS

The Recombinant Clones Are Specifically Homologous to Human Actin mRNA. The RNA that hybridized to the human actin clones was specifically enriched for sequences directing the synthesis of a 42,000-dalton protein which comigrated with authentic rabbit α -actin (Fig. 1). Similar results were observed with the other seven clones isolated; however, three of these clones (λ HRL51, λ HRL54, and λ HRL65) hybridized very poorly to mRNAs isolated from HeLa or human muscle cells (unpublished observations). In each case, the protein products synthesized *in vitro* were further identified as actin by coelectrophoresis on two-dimensional gels with authentic actin standards. A typical example is shown in Fig. 2.

As a further proof that the protein synthesized *in vitro* was actin, the partial proteolysis products of the mRNA-directed translation product selected by one of the clones were compared with those of authentic rabbit α -actin. The products generated by partial V8 protease digestion of the 42,000-dalton putative actin protein comigrated with those of authentic rabbit α -actin (Fig. 3).

Physical Characterization of the Human Actin Gene Clones.

We next compared the DNA structure of the individual clones to assess their genetic identity. Fig. 4A shows the ethidium

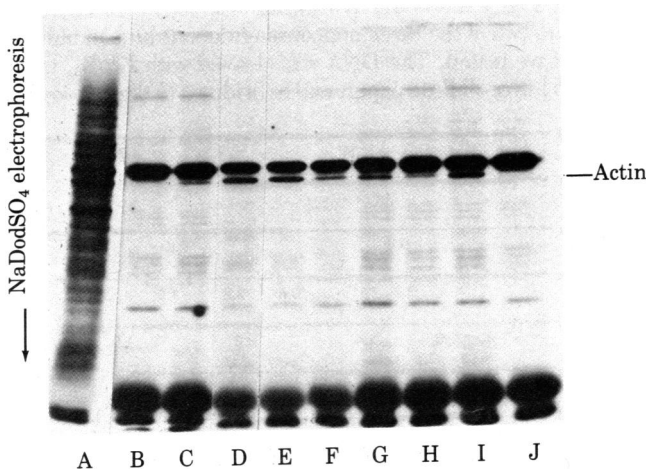


FIG. 1. Fluorogram of radiolabeled proteins synthesized *in vitro* by mRNA that hybridized to seven human genomic actin clones. Total cell RNA was hybridized to nitrocellulose filters to which DNA from each of the actin clones had been bound. The specifically hybridizing mRNA was translated in reticulocyte lysate, and the [³⁵S]methionine-labeled products were analyzed by electrophoresis on a NaDodSO₄/12.5% polyacrylamide gel, followed by fluorography. Lanes: A, HeLa cell RNA; B, mRNA selected by blank filter; C, mRNA, λ HRL84; D, mRNA, λ HRL45; E, mRNA, λ HRL35; F, mRNA, λ HRL34; G, mRNA, λ HRL25; H, mRNA, λ HRL24; I, mRNA, λ HRL21; J, no exogenous mRNA added. The position of the unlabeled α -actin marker detected by Coomassie blue staining is indicated on the right.

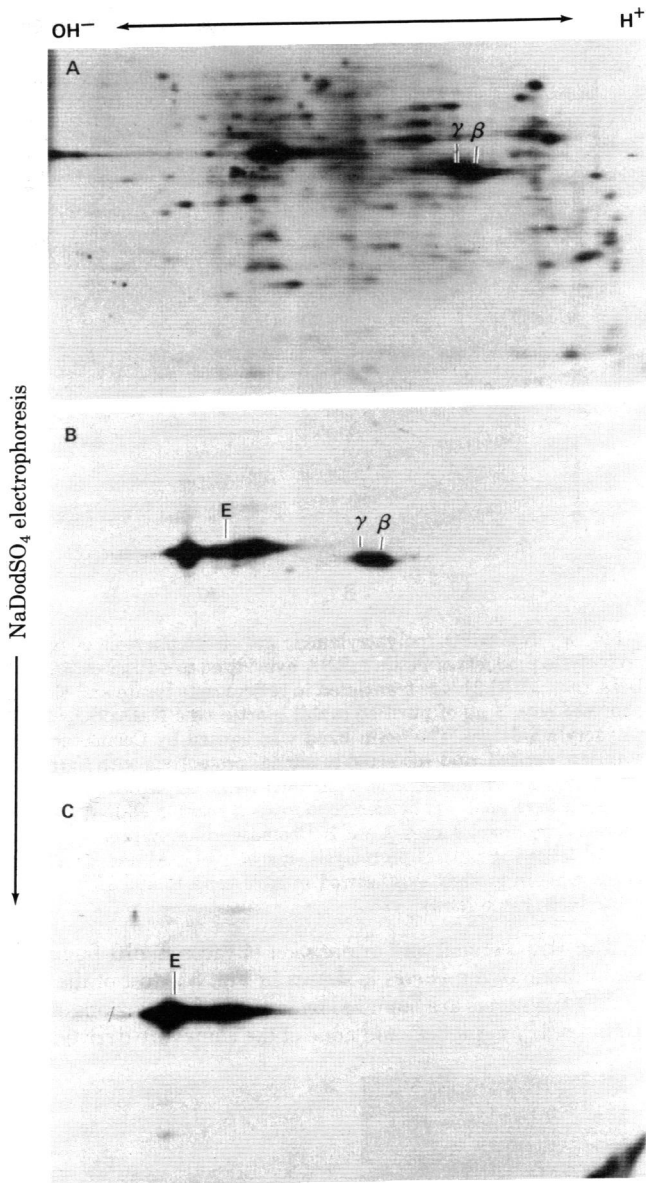


FIG. 2. Fluorogram of two-dimensional gel electrophoresis of *in vitro* translation products. mRNA was prepared and translated as described in the legend of Fig. 1. The products were analyzed by two-dimensional gel electrophoresis as described (22). (A) Total HeLa cell RNA. (B) mRNA selected by λ HRL24. (C) No exogenous RNA added. The identification of β - and γ -actin was by alignment of the radioactive spots with stained nonradioactive actin from rat myoblasts. E, protein synthesized endogenously by the reticulocyte lysate.

bromide stain of the *Eco*RI restriction pattern of the DNA from nine of the recombinant phages after fractionation through an agarose gel. Although a number of bands from different clones appeared to be of almost identical size, additional restriction enzyme analysis demonstrated that none of these human actin clones shared any *Eco*RI fragments (data not shown). Similarly, the three clones with poor homology to human actin mRNA did not share any *Eco*RI fragments with each other or with any of the other clones (data not shown). Because these three clones hybridized so poorly to human actin mRNAs, they were not further studied. Two other clones, λ HRL64 and λ HRL82, were shown to contain identical *Eco*RI fragments to clones λ HRL23 and λ HRL24, respectively. Only a single *Eco*RI fragment in each of these clones hybridized to the heterologous actin gene probes (Fig. 4B).

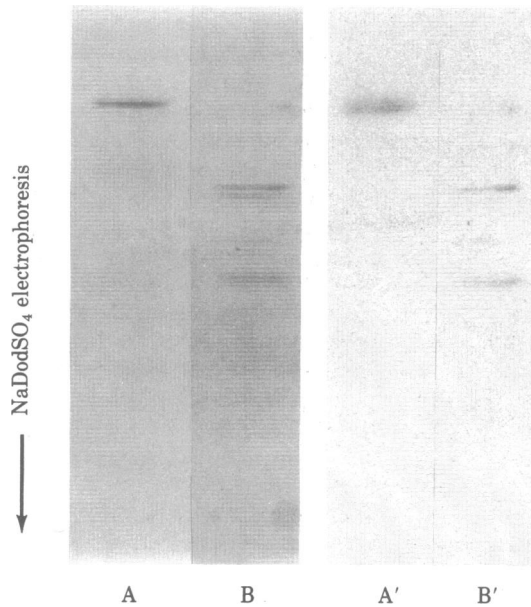


FIG. 3. NaDodSO₄/polyacrylamide gel electrophoresis of partial proteolysis products of actin. mRNA hybridized to a filter containing DNA from λ HRL21 was translated in reticulocyte lysate and electrophoresed with 3 μ g of purified rabbit α -actin on a NaDodSO₄/12.5% polyacrylamide gel. The actin band was located by Coomassie blue staining, excised, and subjected to partial proteolysis with *Staphylococcus aureus* V8 protease as described (24). The partial proteolysis products were analyzed by electrophoresis through a NaDodSO₄/15% polyacrylamide gel. Lanes: A and B, Coomassie blue stain of actin prior to and following partial proteolysis, respectively; A' and B', fluorogram of actin product synthesized *in vitro* prior to and after partial proteolysis, respectively.

The arrangement and orientation of these *Eco*RI fragments within nine of the clones is shown in Fig. 5. Most of the actin coding fragments are bounded by large flanking regions of non-actin-coding sequence, and none of the clones are derived from

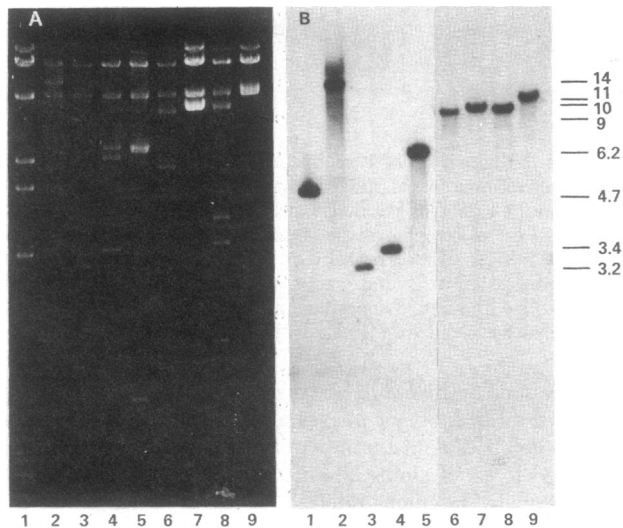


FIG. 4. Restriction enzyme analysis of human actin clones. DNA (1 μ g) from each of the recombinant phages was digested with *Eco*RI and analyzed by electrophoresis on a 0.6% agarose gel. (A) Ethidium bromide stain. (B) Nitrocellulose blot hybridized to ³²P-labeled *Drosophila* cytoplasmic actin clone. Lanes: 1, λ HRL23; 2, λ HRL83; 3, λ HRL45; 4, λ HRL34; 5, λ HRL25; 6, λ HRL35; 7, λ HRL84; 8, λ HRL24; 9, HRL21. Identical hybridization results also were obtained with the other two heterologous actin gene probes. Sizes in kilobases of each of the hybridizing *Eco*RI fragments, indicated on the right, were determined by coelectrophoresis of λ DNA cut with *Hind*III.

overlapping regions of the genome. We also determined the arrangement of the *Eco*RI fragments of clones λ HRL51, λ HRL54, and λ HRL65 (data not shown): The fact that they shared no *Eco*RI fragments in common with each other or with any of the other 9 actin clones and that all the clones had significantly different restriction enzyme sites (unpublished observations) permits us to conclude that we have isolated 12 non-allelic and nonsimilar *Eco*RI fragments containing actin coding sequences that are derived from nonoverlapping regions of the genome.

The Human Genome Contains Actin Coding Regions on at Least 25 Distinct *Eco*RI Fragments. In order to directly assess the number of actin gene fragments contained in the human genome, ³²P-labeled DNA from the chicken β -actin cDNA clone or from each of the two *Drosophila* actin clones was hybridized to nitrocellulose blots of *Eco*RI-digested nuclear DNA from a human individual. The data in Fig. 6 shows that the same 25–30 *Eco*RI fragments hybridized to the chicken β -actin and *Drosophila* muscle actin gene probes. The *Drosophila* cytoplasmic actin gene probe yielded an identical hybridization pattern (data not shown). This result supports the conclusion that the human genome contains actin coding sequences on 25–30 different *Eco*RI fragments. Among the bands that hybridized were fragments that corresponded to the sizes of each of the actin gene encoding *Eco*RI fragments that we have cloned. Ten to 15 bands of hybridization were discernible when the human DNA was cleaved with the enzymes *Bam*HI or *Hind*III (data not shown). This finding might imply that some of the *Eco*RI actin gene-containing fragments are clustered in the genome. However, the fact that most of the actin clones contained coding regions on a single *Eco*RI fragment surrounded by non-actin-encoding *Eco*RI fragments suggests that the *Eco*RI actin gene fragments are separated on the chromosome by more than 10 kilobases.

Polymorphism in the Human Actin Gene Family. To search for population polymorphism in actin genes, we examined DNA isolated from different individuals for changes in actin gene sequences that would result in alterations of restriction fragment patterns. Fig. 7 shows the hybridization spectrum to DNA isolated from two of the three nonconsanguineous human individuals that we tested. The DNA was cleaved with *Eco*RI, transferred to nitrocellulose paper, and hybridized to the *Drosophila*

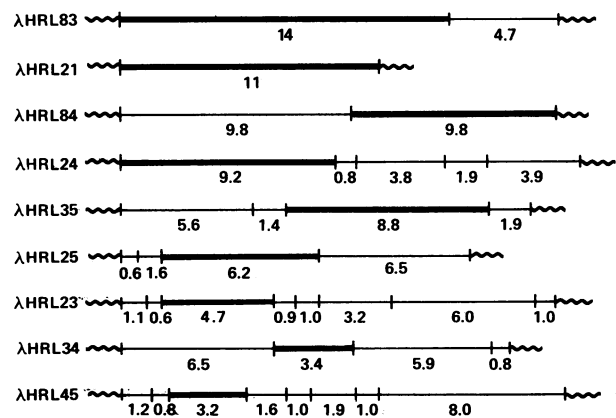


FIG. 5. *Eco*RI restriction endonuclease maps of recombinant phages containing actin gene coding regions. The *Eco*RI fragments are delineated by vertical lines, and the fragments containing actin coding sequences are shown by the heavy black lines. The Charon 4A arms are indicated by squiggly lines and the sizes of the *Eco*RI fragments are indicated beneath each fragment. The maps were deduced by partial and double restriction enzyme digests in conjunction with hybridization of appropriate nick-translated subcloned fragments to nitrocellulose blots of these digests.

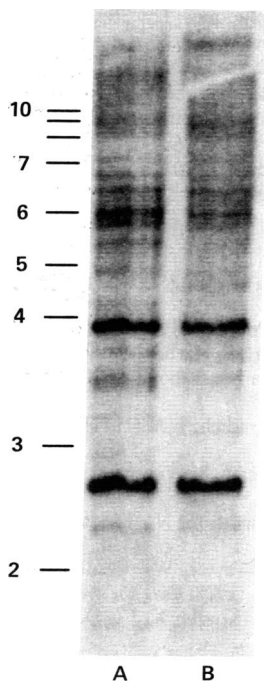


FIG. 6. Determination of the number of *EcoRI* fragments in the human genome homologous to actin genes. DNA (2 μ g) from an individual was digested with *EcoRI*, electrophoresed through a 0.6% agarose gel, transferred to a nitrocellulose filter, and hybridized with a nick-translated probe prepared from the *Drosophila* muscle actin gene clone (lane A) or from the chicken β -actin cDNA clone (lane B). Size of the DNA in kilobases is indicated on the left.

cytoplasmic actin gene probe. The pattern of hybridizing bands was identical for each of the three DNAs examined. Similarly, when DNA from these three individuals was cleaved with the restriction enzymes *Bam*HI or *Hind*III and hybridized to any of the three heterologous probes, no differences in the hybridization banding patterns were observed between the individual's DNA for either of the enzymes utilized (data not shown). Thus, at the level of restriction enzyme site and fragment length changes, we were unable to detect any polymorphism in actin gene-containing fragments in this small population sample.

DISCUSSION

We have described the isolation of 12 different nonoverlapping recombinant phages containing human actin coding sequences. These recombinants were isolated from a library of cloned human DNA fragments derived from a single individual. Three of these clones, originally selected by hybridization to the heterologous actin probes, hybridized poorly to human actin mRNA. These three human clones may contain actin gene sequences that have diverged from the major classes of transcribed human actin genes. The observation that they were selected by their ability to hybridize to chicken and *Drosophila* actin gene probes may be explained by the fact that the DNA-DNA hybridizations were performed under less stringent conditions than were the RNA-DNA hybridizations.

We conclude that there are many regions in the human genome that contain actin coding sequences. Between 25 and 30 distinct *EcoRI* fragments on genomic Southern blots hybridize to chicken or *Drosophila* actin gene probes. Likewise, 10-15 bands hybridized with these probes after *Bam*HI or *Hind*III digestion. Cleveland *et al.* (15) reported observing at least six human *Hind*III fragments that hybridize to the same chicken β -actin cDNA clone. Examination of their autoradiograms sug-

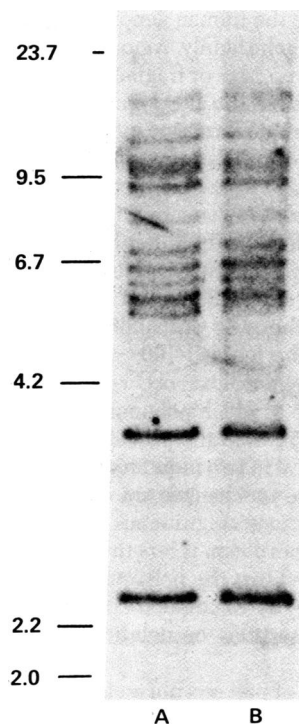


FIG. 7. Search for polymorphism in human DNA restriction fragments containing actin coding sequences. DNA (2 μ g) from two different individuals (lanes A and B) were digested with *EcoRI*, subjected to electrophoresis through a 0.6% agarose gel, transferred to nitrocellulose filter, and hybridized with a nick-translated probe prepared from the *Drosophila* cytoplasmic actin clone. Length of DNA is shown in kilobases.

gests that there may, in fact, be more than 10 bands of hybridization. The specificity of the heterologous probes for human actin sequences is demonstrated by the observation that human actin gene probes, derived from the clones that we have isolated, hybridize to the same set of fragments. Furthermore, no clones that hybridize to the heterologous probes fail to hybridize to human actin mRNA. Finally, calculations based on the probability of randomly selecting 12 out of 14 nonidentical clones or of randomly selecting 2 out of 14 identical clones indicate that there is a 98% chance[†] that over 20 unique *EcoRI* actin coding fragments exist within an individual genome. To determine whether these 20 or more different fragments each contain an entire actin gene, the coding regions of each of the cloned fragments will have to be located. Our preliminary results with the 12 clones suggests that each contains at least most of an actin gene.

We have utilized the chicken and *Drosophila* actin gene

[†] The probability was calculated as follows: It is assumed that there was an infinite number of copies of each clone present in the phage pool from which an aliquot was selected and analyzed. If there are n unique, equally represented actin-coding *EcoRI* fragments, the probability of observing 12 different fragments and 2 duplicates of 2 of the different fragments is given by

$$p(n) = \frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{n-11}{n} \cdot \frac{12}{n} \cdot \frac{11}{n} = \frac{132n!}{(n-12)!n^{14}}$$

This probability was determined for different values of n and was found to be highest for $30 < n < 60$. The probability that there exists more than 20 actin-coding *EcoRI* fragments is given by

$$1 - \left(\sum_{n=12}^{n=20} p(n) / \sum_{n=12}^{n=\infty} p(n) \right)$$

probes to examine the human genome for population polymorphism in the actin gene family. Among three different individual DNAs examined, no base or fragment length changes were detected that resulted in the alteration of restriction enzyme sites for the three different restriction enzymes tested. Although this result describes only a small population sample, this assay should, in theory, be able to detect even less than a 1% rate of polymorphism between any two individuals. This sensitivity derives from the fact that each *EcoRI* restriction fragment compares the identity of 12 nucleotides between two individuals. Because we can detect 25 *EcoRI* fragments on the genomic Southern blots, then the probability of two individuals being polymorphic at the loci is $[(100 - N)^{12}]^{25}$, in which N is the percentage of polymorphism occurring in the population. If, for example, the rate at which polymorphism occurs in the human genome is 1% (25), then the probability that all 25 *EcoRI* fragments are identical in two individuals is less than 5%. Our finding that there are no changes among 25 *EcoRI* fragments between three nonconsanguineous individuals suggests that population polymorphism is less than 1% at the actin gene loci. It should be noted that the polymorphism level may be far less than 1% because our calculation does not take into account the possibility of insertions or deletions occurring between the *EcoRI* sites.

The low level of polymorphism in the human actin gene loci contrasts the results of similar types of experiments which examined other human gene loci for population polymorphism. Jeffreys (25) estimated that population polymorphism was occurring at a 1% rate in the intervening sequences of human β - and γ -globin gene loci but at a lower rate in the coding sequences. Others (26), using a randomly cloned single copy DNA probe, found a high frequency of sequence variation at this arbitrary locus. Thus, it would appear that selective pressure prevents random base changes from being maintained in the actin gene sequences or in their immediate flanking regions. This unexpected conservation of actin gene nucleotide sequences may be a consequence of the constraints imposed on the sequence of a structural protein which interacts with many cell components. Because the actin coding sequences do not span all of the *EcoRI* sites, this conservation of primary DNA sequence outside actin coding regions might reflect the importance of RNA or DNA secondary structure in regions involved with transcriptional or translational regulation.

We would like to thank Drs. Don Cleveland and Marc Kirschner, Sally Tobin and Brian McCarthy, and Eric Fyrberg and Norman Davidson for providing us with their actin clones; Dr. Tom Maniatis for generously donating the human library and the human fetal DNA used for constructing the library; Dr. James Spudich for the rabbit α -actin standard; Drs. Gary Peltz, Susan Guttman, and Helen Blau for familiarizing

us with protein gel techniques; Dr. Steven Embury for his kind gift of human DNA; Dr. Doug Wallace for guidance with growing the HeLa cells; and Drs. Rob Maxson, Geoffrey Childs, and Ron Cohn for their helpful suggestions and comments. This work was supported in part by grants from the National Institutes of Health, the American Cancer Society, and by the Veterans Administration. L.K. is an Investigator, Howard Hughes Medical Institute. J.N.E. is a trainee of the Medical Scientist Training Program.

- Pollard, T. D. & Wehing, R. R. (1974) *CRC Crit. Rev. Biochem.* **2**, 1-65.
- Goldman, R., Pollard, T. & Rosenbaum, J., eds. (1976) *Cell Motility*, Cold Spring Harbor Conferences on Cell Proliferation (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 3.
- Clarke, M. & Spudich, J. A. (1977) *Annu. Rev. Biochem.* **46**, 797-822.
- Vandekerckhove, J. & Weber, K. (1978) *J. Mol. Biol.* **126**, 783-802.
- Bender, W., Davidson, N., Kindle, K., Taylor, W., Silverman, M. & Firtel, R. (1978) *Cell* **15**, 779-788.
- McKeown, M., Taylor, W., Kindle, K., Firtel, R., Bender, W. & Davidson, N. (1978) *Cell* **15**, 789-800.
- Fyrberg, E. A., Kindle, K. L., Davidson, N. & Sodja, A. (1980) *Cell* **19**, 365-378.
- Tobin, S. L., Zulauf, E., Sanchez, F., Craig, E. A. & McCarthy, B. J. (1980) *Cell* **19**, 121-131.
- Durica, D. S., Schloss, J. A. & Crain, W. R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5683-5687.
- Palmiter, R. D. (1974) *Biochemistry* **13**, 3606-3615.
- Childs, G., Maxson, R. & Kedes, L. H. (1979) *Dev. Biol.* **73**, 153-173.
- Sures, I., Maxam, A., Cohn, R. H. & Kedes, L. H. (1976) *Cell* **9**, 495-502.
- Kaiser, A. D. & Hogness, D. S. (1960) *J. Mol. Biol.* **2**, 392-415.
- Kan, Y. W., Dozy, A. M., Trecartin, R. & Todd, D. (1977) *N. Engl. J. Med.* **297**, 1081-1084.
- Cleveland, D. W., Lopata, M. A., MacDonald, R. J., Cowan, N. J., Rutter, W. J. & Kirschner, M. W. (1980) *Cell* **20**, 95-105.
- Rigby, P. W., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237-251.
- Maniatis, T., Hardison, R., Lacy, E., Lauer, J., O'Connell, C., Quan, D., Sim, G. & Efstratiadis, A. (1978) *Cell* **15**, 687-701.
- Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180-182.
- Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503-517.
- Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* **67**, 247-256.
- Laemmli, U. K. (1970) *Nature (London) New Biol.* **227**, 680-685.
- O'Farrell, P. H. (1975) *J. Biol. Chem.* **250**, 4007-4021.
- Lasky, R. A. & Mills, A. D. (1975) *Eur. J. Biochem.* **56**, 335-341.
- Cleveland, D., Fischer, S., Kirschner, M. & Laemmli, U. (1977) *J. Biol. Chem.* **252**, 1102-1106.
- Jeffreys, A. J. (1979) *Cell* **18**, 1-10.
- Wyman, A. R. & White, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6754-6758.