Proc. Natl. Acad. Sci. USA
Vol. 77, No. 3, pp. 1351-1355, March 1980
Biochemistry

Structure of the α-fetoprotein gene in the mouse

(genomic library/cloning/restriction mapping/electron microscopy/intervening sequences)

MICHAEL B. GORIN AND SHIRLEY M. TILGHMAN

The Institute for Cancer Research, The Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111

Communicated by Robert Palese Perry, December 10, 1979

ABSTRACT **The mouse α-fetoprotein mRNA is the product of a single-copy gene whose mRNA coding sequences are represented discontinuously in the genome. Several *Eco*RI genomic fragments which contain portions of the α-fetoprotein gene have been cloned by using the EK2 vector λgt*WES*·λB. In addition, a mouse genomic library has been screened to obtain a 15.75-kilobase segment of DNA that includes more than 85% of the α-fetoprotein coding sequence. Analyses by restriction endonuclease mapping and electron microscopy showed that the mRNA sequence is interrupted by at least 11 intervening sequences which occupy 90% of the cloned DNA.**

α-Fetoprotein (AFP) is a 70,000-dalton protein that is synthesized in the yolk sac and liver of all mammals during embryonic development and is secreted into the amniotic fluid and blood (1–4). After birth, the rate of AFP synthesis in liver decreases until the serum AFP level reaches a basal level of 0.01% of that in fetal serum. Reinitiation of AFP synthesis in the adult is observed upon neoplastic transformation of or injury to the liver and in teratocarcinomas (5, 6).

Hybridization analyses have shown that alterations in AFP synthesis are the result of changes in the amount of AFP mRNA in liver cells (7, 8). As a step toward elucidating the basis of AFP gene regulation, we have examined the structure of the AFP gene in the mouse. Cloned AFP cDNA probes, termed pBR322·AFP1 and AFP2 and described by Tilghman *et al.* (9), were used to detect the presence of the single-copy AFP gene in mouse genomic DNA and to show that its coding sequences were represented discontinuously in at least three *Eco*RI fragments (9). In this paper, we describe the structural organization of the coding sequences of the AFP gene as revealed by the study of a cloned 15.75-kilobase (kb) fragment of mouse genomic DNA.

MATERIALS AND METHODS

Large molecular weight DNA, extracted from Swiss mouse embryos (10), was cleaved with *Eco*RI, chromatographed on an RPC-5 column, and analyzed for fragments containing AFP-coding sequences as described (9, 11). DNA (3 μg) of selected pooled fractions was ligated to 7 μg of purified λgt*WES*·λB outer *Eco*RI fragments (11) in a final volume of 100 μl containing 1 mM ATP, 10 mM $MgCl_2$, 20 mM Tris·HCl (pH 7.5), and 3 units of T4 DNA ligase (New England BioLabs); incubation was at 4°C for 24 hr. Then, additional ligase (1 unit) and ATP (10 nmol) were added and the incubation was continued at 4°C for 16–24 hr. The ligation mixture was heated at 65°C for 5 min, and 5-μl aliquots were used in individual *in vitro* packaging reactions (12). The DNase-treated mixture was plated onto LB plates (1.0% tryptone, 0.5% yeast extract, 0.5% NaCl, 1.5% agar) with the host, *Escherichia coli* LE392 (11).

Fifty petri dishes with 5000–12,000 phage plaques were screened by the procedure of Benton and Davis (13), with DNA fragments derived from restriction endonuclease digests of the chimeric AFP cDNA plasmid pBR322·AFP1 or pBR322·AFP2 (9) that had been nick-translated to 4–10 × 10^7 cpm/μg in the presence of [^{32}P]dCTP (14). Regions on the LB plates that contained a hybridizing plaque, identified by comparison to the replica nitrocellulose filters, were scraped off, resuspended in 50 mM Tris·HCl, pH 7.5/10 mM $MgCl_2$, and replated at low density in order to isolate individual positive phage plaques. Selected phage were grown in LB medium with *E. coli* strain DP50*sup*F as the host (15).

A mouse genomic library, prepared from random partial *Hae*III/*Alu* I digests of BALB/c embryonic DNA and the outer *Eco*RI arms of the vector Charon 4A (16, 17), was provided by J. G. Seidman. The library was plated to a density of 20,000 phage per plate and screened as described above.

Restriction endonuclease mapping of phage DNA was performed by using conditions recommended by the supplier except that albumin was omitted. Restriction fragments were analyzed by agarose and acrylamide (18) gel electrophoresis and stained with ethidium bromide (25 μg/ml). The restricted DNA was transferred from agarose gels to nitrocellulose filters (Millipore) by the method of Southern (19) except that the elution buffer was ×6 standard saline citrate (NaCl/NaCit). Pretreatment, hybridization, and washing of the filters were performed as described by Wahl *et al.* (20). Filters were exposed at −70°C to Kodak XR-1 backed by DuPont Lightning Plus intensifying screens.

The genomic fragments were subcloned into the plasmid vector pBR322 after cleavage of the phage DNA with *Eco*RI and ligation into the *Eco*RI site of pBR322 (9). The *E. coli* strain χ1776 was transformed with the ligated plasmid DNA as described (9). Colonies were screened by the method of Grunstein and Hogness (21) with labeled restriction fragments from chimeric AFP cDNA plasmids.

Heteroduplexes between pBR322·AFP1 and the subcloned *Eco*RI genomic fragments were formed by cleaving the plasmids with *Sal* I, extracting with phenol, and precipitating the linearized DNA with 2 vol of ethanol at −20°C. The DNAs were mixed at a 1:1 molar ratio, denatured in 0.3 M NaOH for 15 min at 37°C, and, after neutralization, allowed to renature for 1 hr at room temperature in 0.1 M Tris·HCl, pH 7.5/10 mM EDTA at a DNA concentration of 5–10 μg/ml.

Duplexes between *Eco* B, an internal *Eco*RI fragment of the genomic clone, and AFP mRNA, purified from yolk sacs of 18-day embryos (9), were formed by incubating *Eco*RI-cleaved plasmid DNA (50 μg/ml) and AFP mRNA (50 μg/ml) in 70% formamide/0.5 M NaCl/0.1 M *N*-tris(hydroxymethyl)methylglycine pH 8.0/10 mM EDTA at 52°C for 16 hr (22).

The hybrids were diluted to 2–4 μg of DNA per ml and spread for electron microscopy as described (11). Molecules were viewed in a Philips 300 microscope at ×16,000 magnifi-

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.

Abbreviations: AFP, α-fetoprotein; kb, kilobase(s); NaCl/NaCit, 0.15 M NaCl/0.015 M Na citrate; bp, base pair(s).

cation and measured with a Numonics 1224 digitizer. The distances between the *Hin*dIII and *Sal* I sites in pBR322, as determined by Sutcliffe (23), were used as internal standards for the calculation of base pairs in the heteroduplexes in Fig. 3. For the R-D loops in Fig. 4, simian virus 40 nicked circles were included in the spreading solution.

The experiments that used recombinant DNA were performed in accordance with the National Institutes of Health *Guidelines for Research Involving Recombinant DNA Molecules.*

RESULTS AND DISCUSSION

Cloning of AFP-Encoding Mouse Genomic DNA. We initially approached the isolation and cloning of the AFP gene by using *Eco*RI fragments that had been enriched approximately 10-fold by RPC-5 chromatography (9). At least three fragments 6.0, 3.0, and 5.2 kb long had been identified as containing AFP coding sequences by hybridization to AFP structural gene probe. These corresponded to 5′, middle, and 3′ coding regions, respectively. Selected pools of DNA containing these fragments were ligated to the isolated outer *Eco*RI fragments of λgt*WES*λB (11) and the concatameric DNA was packaged into viable phage particles by the procedure of Sternberg *et al.* (12). Three recombinant phage were subsequently detected by hybridization (13), isolated, and characterized as containing AFP coding sequences. These phage, designated λgt*WES*·AFP4, λgt*WES*·AFP5, and λgt*WES*·AFP6, contained *Eco*RI fragments 1.48, 2.75, and 4.75 kb long. The last two fragments corresponded to the 3.0- and 5.2-kb fragments originally detected on the Southern hybridization blot. The 1.48-kb insert was unexpected and was found by hybridization to contain a small segment of AFP sequence derived from the 5′ end of the mRNA which placed it immediately adjacent to the two other fragments (data not shown). Our inability to detect this fragment on the original genomic blots was the result of the sensitivity of the system used. In later experiments, when probes of higher specific activity were used along with a more efficient hybridization procedure (20) we were able to detect this fragment (data not shown). However, this experience pointed out one of the obvious limitations in genomic cloning of large genes with limit digests of *Eco*RI fragments.

Recently, a mouse genomic library, constructed by Jon Seidman, was generously provided to us, and we have used it to clone those regions of the AFP gene that were not selected in the *Eco*RI fragment screen. By this approach, we hoped to overcome the additional problem of missing regions of the gene that might contain no coding sequences. In addition, we hoped to obtain both 5′ and 3′ flanking sequences. Approximately 10^6 phage particles of the library were screened by hybridization, and three positive phage were identified. One of these, λCh4·AFP7, which contained >80% of the AFP coding sequences, was chosen for detailed characterization. The two other phages contained the 3′ end of the gene itself, along with 4–8 kb of 3′ flanking sequences.

Restriction Endonuclease Mapping of λCh4·AFP7. When λCh4·AFP7 DNA was cleaved with *Eco*RI, six internal fragments were generated, labeled A–F in Fig. 1. That the band at 4.75 kb constituted a doublet was demonstrated by cleavage of the *Eco*RI digest with *Xho* I, which does not cleave the outside λCh4A arms, and cleaves only the *Eco*B fragment within the mouse DNA (Fig. 1, lane b). In this way, each of the *Eco*RI fragments could be separated on agarose gels for hybridization to specific probes derived from different regions of the AFP mRNA, shown diagrammatically in Fig. 1.

The orientation of those *Eco*RI fragments derived from the 5′ end of the mRNA were determined by using probes 1, 2, and

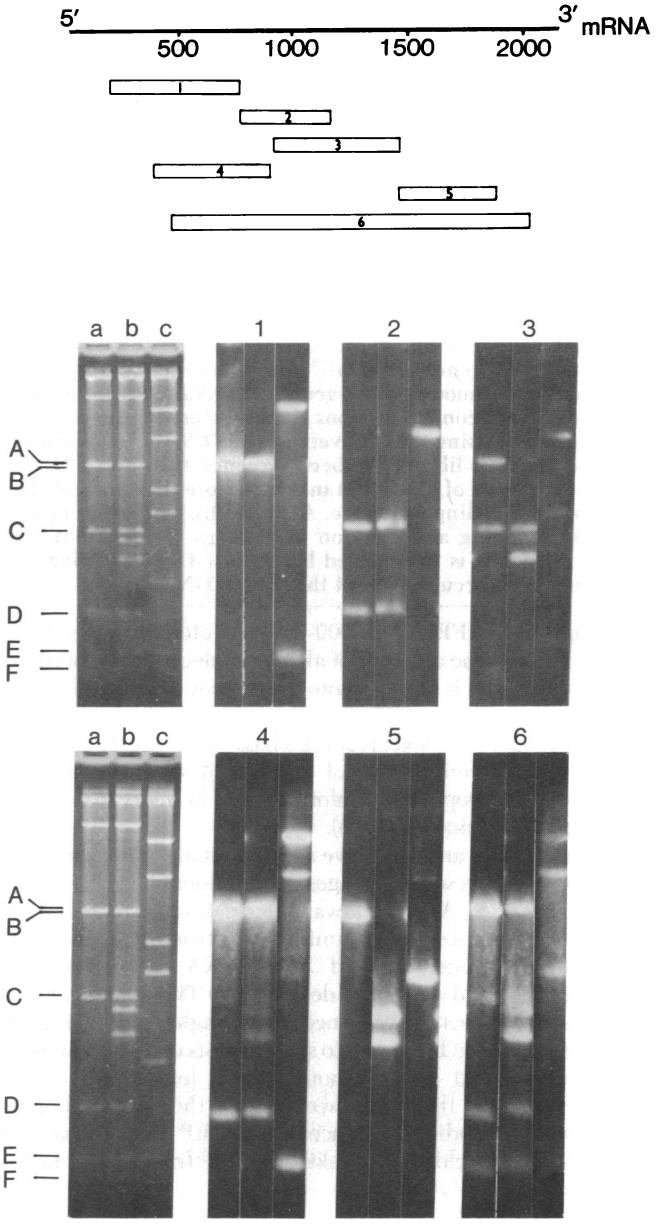

FIG. 1. Orientation of *Eco*RI fragments within λCh4·AFP7. Specific restriction fragments, identified in the numbered boxes of the diagram (*Top*), were prepared from pBR322·AFP1 and pBR322·AFP2 (probes 1–5) or pBR322·AFP3 (probe 6). The regions of AFP mRNA from which they were derived are indicated. In the gels, 0.5 μg of aliquots of λCh4·AFP7 DNA were digested with *Eco*RI (lane a), *Eco*RI/*Xho* I (lane b), or *Sac* I (lane c) and electrophoresed on replica 1% agarose gels. Ethidium bromide staining of the gels are shown at the left. Sizes of restriction endonuclease fragments were determined by comparison to fragments of a *Hin*dIII digestion of λ*cI*857 and a *Hae* III digestion of ColE1. The *Eco*RI fragments, lettered A–F, are 4.75, 4.75, 2.75, 1.48, 1.1, and 0.95 kb long. In panels 1–6 are hybridizations of the digests of λCh4·AFP7 DNA with probes 1–6, respectively.

4 (Fig. 1). Probe 1 recognized only *Eco* A, which placed it at the farthest 5′ end of the gene. Probe 2 detected *Eco* D and *Eco* C but not *Eco* A. Probe 4, which overlaps probes 1 and 2, was used to orient *Eco* D and *Eco* C. It hybridized to *Eco* A and *Eco* D but failed to detect *Eco* C. Thus, the relative order of these fragments was *Eco* A–D–C.

The hybridization of probe 3 to both *Eco* C and *Eco* B in the *Eco*RI/Xho digest established that *Eco* B was adjacent to *Eco* C. This was confirmed by using probe 5, which hybridized only

to *Eco* B. The farthest 3′ coding fragment, *Eco* E, was identified by its hybridization to probe 6, which contained all but approximately 50 base pairs (bp) of the 3′ end of the AFP cDNA (unpublished results). The smallest *Eco*RI fragment, *Eco* F, contained no AFP coding sequence and was mapped to the 3′ side of *Eco* E by calculating the distances between restriction endonuclease sites in the outer arms of λCh4A and the nearest internal sites that could be mapped by hybridization analysis.

To confirm and to extend the restriction map of λCh4·AFP7, similar experiments were performed with λCh4·AFP7 DNA and other restriction endonucleases. Only one of these, a digestion with *Sac* I, is shown in Fig. 1, lane c. Thereby the map shown in Fig. 2 was derived. The AFP genomic sequences in λCh4·AFP7 are drawn opposite to the normal λ left-to-right convention, as is indicated, because the 5′ end of the AFP sequences were adjacent to the right arm of λCh4A.

The broken lines connecting the known restriction sites in AFP cDNA (9) and the corresponding genomic sites in Fig. 2 were determined from experiments like those shown in Fig. 1. For example, in a *Bam*HI digest of λCh4·AFP7, only one fragment hybridized to probe 1, which was generated by a *Bam*HI cleavage site in pBR322·AFP1 (see Fig. 1). As would be predicted, the adjacent *Bam*HI fragment in the genome hybridized to probe 2. In this way, we were able to align those sites that were derived from AFP coding sequences. Confirmation of these results was obtained by electron microscopy.

Identification of the AFP Coding Sequences in λCh4·APF7. In order to identify the map positions of the AFP coding regions in λCh4·AFP7, the *Eco*RI fragments were individually subcloned into the *Eco*RI site of the plasmid pBR322 as described (9). The chimeric plasmids generated were identified by hybridization. Restriction endonuclease mapping was used to confirm their identity and to determine their 5′-to-3′ orientation within the plasmid vector.

To identify the coding segments in *Eco* A, D, and C, heteroduplexes were formed between each genomic plasmid and pBR322·AFP1, an AFP cDNA clone containing a 960-bp insert in the *Hin*dIII site of pBR322 that includes the 5′ mRNA sequences as represented by probes 1 and 2 (Fig. 1) (9). The 29-bp difference between the *Eco*RI and *Hin*dIII sites of insertion (23) did not affect the stability of the duplexes and was subtracted from subsequent measured values. Each plasmid was first linearized with *Sal* I, which does not cleave within the inserted DNAs but does cut pBR322 DNA at a site 650 bp from the *Eco*RI site (23). In this way, the orientation of the heteroduplex structures visualized in the electron microscope could be immediately determined. In addition, the pBR322 homoduplex "arms" provided internal molecular weight markers for each molecule.

Fig. 3*a* shows a heteroduplex between *Eco* A and AFP1 along with its 5′-to-3′ orientation. The presence of four loops of single-stranded DNA—two outside "substitution" loops which represent intervening sequence DNA in the genomic clone on one side and coding sequence in the cDNA clone on the other side, and two internal "deletion" loops which mark the positions of genomic intervening sequences. These are connected by three separate homologous regions, representing three coding sequences in *Eco* A. Measurements of a series of these molecules produced the results shown in Fig. 2. The presence of single-stranded AFP1 DNA at the 5′ end of the heteroduplex shows that *Eco* A must not contain the 5′ end of the gene. We previously determined (9) that AFP1 itself is missing 100–150 bp of 5′ AFP mRNA sequence. Because the AFP1 side of the substitution loop was measured to be 150 ± 30 bp, we conclude that approximately 250 bp of AFP coding sequence is unaccounted for in λCh4·AFP7.

Heteroduplexes between AFP1 and *Eco* D and *Eco* C are shown in Fig. 3 *b* and *c*, respectively. In each case, only one homologous region was observed, surrounded by two substitution loops. In the case of *Eco* C, the absence of measurable single-stranded cDNA-derived DNA in the 3′ loop indicates that *Eco* C contains sequences very close to but not including the *Hin*dIII site which defines the 3′ end of AFP1. The results of the measurements of these molecules are shown in Fig. 2.

The distribution of coding sequences in *Eco* B was determined by the formation of heteroduplexes between the isolated *Eco* B fragment and AFP mRNA. It was not possible to form DNA:DNA duplexes as in Fig. 3 because of the 3′-derived cDNA clones were inserted into the *Pst* I site of pBR322, and stable heteroduplexes could not be obtained. An example of the RNA·DNA duplex is shown in Fig. 4. Five coding segments could be identified, which were interrupted by four single-stranded DNA loops, indicating the presence of intervening

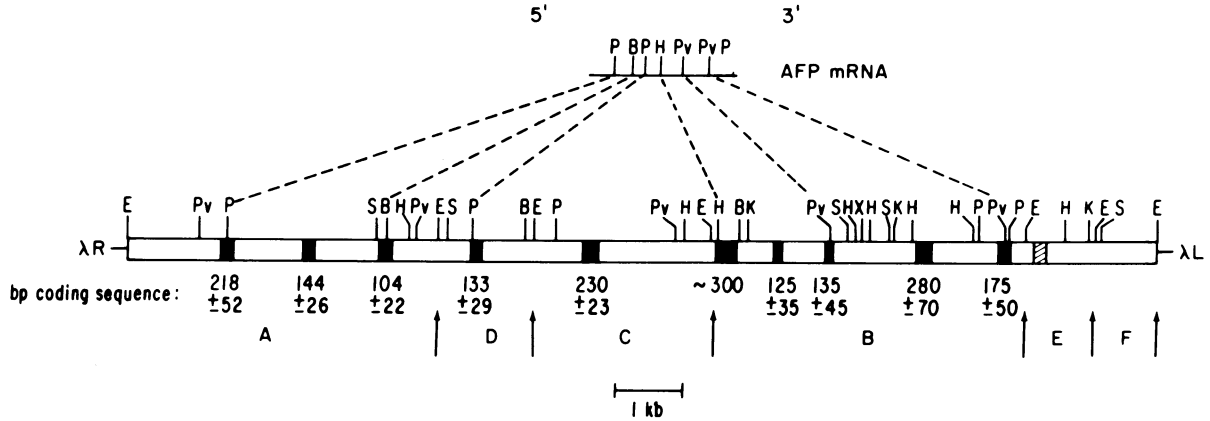

FIG. 2. Restriction endonuclease map of λCh4·AFP7. Top line represents the position of restriction sites in the full-length AFP cDNA, as determined previously (9). The second line illustrates the restriction endonuclease map of λCh4·AFP7, with broken lines connecting the corresponding sites in the mRNA sequence. The open boxes indicate the positions of intervening sequences and the solid boxes indicate the positions of coding regions. The hatched bar represents a region of coding sequence whose exact size and location within the *Eco* E fragment are not established (see text). The sizes and positions of the coding sequences, indicated below the line, were determined by measurements of at least 20 of each of the molecules illustrated in Figs. 3 and 4. The letters below indicate the positions of the *Eco*RI fragments labeled as shown in Fig. 1. The genomic fragment in AFP7 is drawn opposite to its orientation in λCh4A, as indicated by the symbols λR nd λL in the second line. Endonucleases: E, *Eco*RI; B, *Bam*HI; P, *Pst* I; H, *Hin*dIII; Pv, *Pvu* II; S, *Sac* I; X, *Xho* I; K, *Kpn* I.

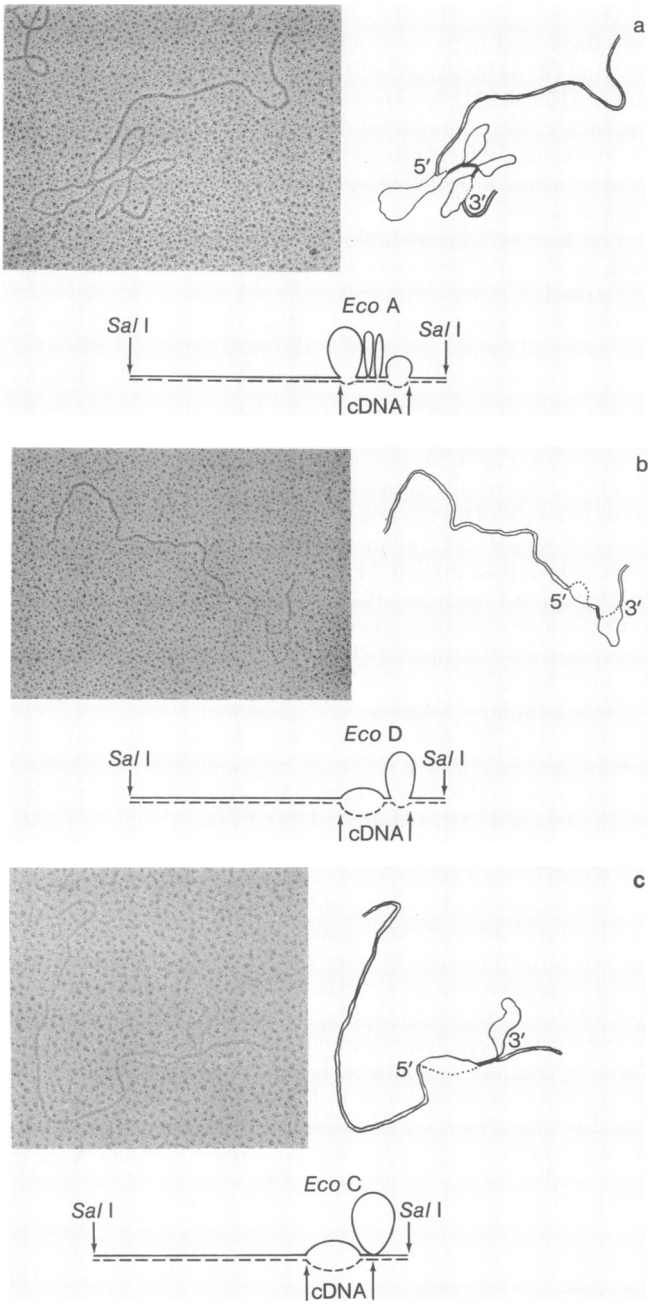

FIG. 3. Electron microscopy of heteroduplexes between pBR322·AFP1 and the AFP genomic fragments. All the plasmids were linearized by cleavage with *Sal* I. The designations of 5′ and 3′ ends of the hybrid molecules correspond to the 5′ and 3′ ends of the AFP mRNA. In the drawings beside the micrographs, the double lines represent the 3.71- and 0.62-kb homoduplexes of pBR322 DNA. Single narrow lines indicate single-stranded genomic DNA, and dotted lines are the single-stranded cDNA segments in pBR322·AFP1. Heavy lines represent the regions of homology between the AFP cDNA and the AFP genomic DNA. The molecules are magnified ×80,000. In the interpretative drawings below, pBR322-genomic DNA is represented by a solid line and pBR322-cDNA, by a dotted line. The two arrows below the drawing designate the borders between pBR322 DNA and the inserts. Heteroduplexes: (*a*) between pBR322-*Eco* A and pBR322-AFP1; (*b*) between pBR322-*Eco* D and pBR322-AFP1; (*c*) between pBR322-*Eco* C and pBR322-AFP1.

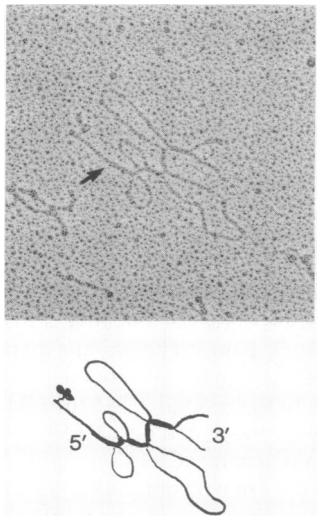

FIG. 4. Electron microscopy of hybrids formed between AFP mRNA and *Eco* B. The genomic insert of pBR322·*Eco*B was excised with *Eco*RI and hybridized to 18S yolk sac RNA. The 5′ and 3′ ends of the mRNA are noted on the drawing. The arrow indicates the approximate end of the homology between the RNA and DNA, which is immediately adjacent to the 5′ end of the DNA fragment.

sequences. The 5′-to-3′ orientation of these loops was determined by performing the experiment with *Sal* I-digested pBR322·*Eco*B (data not shown). The sizes and positions of the coding segments in *Eco* B were determined except for the first which was located at the 5′ end of the fragment. In order to measure this region, it was necessary to discriminate the transition point from single-stranded RNA to RNA·DNA hybrids in each molecule, indicated by the arrow in Fig. 4. This problem arose because the beginning of the coding segment in *Eco* B is less than 50 bp away from the *Eco*RI site (see Fig. 2). Thus, the first coding distance is an estimate only.

The coding region in *Eco* E must be very small because it was only detected faintly with probe 6. In addition, we have been unable to visualize any heteroduplexes between RNA and this fragment. However, we are reasonably sure that it contains the 3′ end of the mRNA sequence, which must be located to the left of the *Hin*dIII site in the fragment. The other 3′ AFP genomic clones isolated, which have substantial 3′ flanking sequences, do not contain further AFP coding segments as determined by hybridization to AFP cDNA (data not shown). Only DNA sequence determination will allow precise identification as the 3′ end of the coding sequences of the gene. These overlapping clones also clearly establish that the *Eco*RI junction between *Eco* F and the left arm of λCh4A is the result of a synthetic *Eco*RI linker because the sequences in *Eco* F are found in an internal 2.45-kb fragment in these 3′ clones. In a similar manner, the *Eco*RI border between *Eco* A and the right λCh4A arm is also generated during the construction of the library, because the AFP coding sequences found in the 4.75-kb *Eco* A fragment are located in a natural 6.0-kb fragment in the BALB/c genome (9).

Structure of the AFP Gene. It is striking that, of the 15.75 kb of genomic DNA in λCh4·AFP7, almost 90% is accounted for by intervening sequences. The 11 coding segments of DNA, none greater than 300 bp long, are interrupted by 11 intervening sequences. This degree of complexity in sequence organization is a minimum estimate because λCh4·AFP7 does not contain the entire AFP gene. As illustrated in Fig. 3*a*, it is missing approximately 250 bp of 5′ coding sequence. We have now screened this mouse embryonic library three times for recombinants that would contain sequences to the 5′ side of λCh4·AFP7 and must conclude that they are absent from the repertoire. Consequently, other mouse genomic libraries must be screened in order to obtain these sequences.

The original *Eco*RI-derived genomic clones, λgt*WES*·AFP4, -5, and -6, were generated from Swiss mouse DNA. By restriction endonuclease mapping and electron microscopy, they have been shown to correspond exactly to the BALB/c-derived *Eco* D, *Eco* C, and *Eco* B fragments, respectively. Thus, at this

macroscopic level, there does not appear to be extensive divergence in the AFP gene between these two strains of mice.

Of the several eukaryotic genes that have been characterized in detail, the AFP gene most closely resembles the rat albumin gene in the degree of complexity of its structure. Sargent *et al.* (24) have recently reported that this gene is split into 14 coding segments over a distance of 15 kb and that each of these segments is less than 300 bp long. We have recently cloned the mouse albumin gene, from RPC-5 column-enriched DNA and the mouse embryonic DNA library, and our preliminary results (unpublished data) suggest that the albumin gene of the mouse closely resembles that of the rat.

There are several compelling pieces of evidence to suggest that albumin and AFP were derived from a common ancestral gene. These include amino acid sequence homology (25) and immunological cross reactivity of the denatured peptides (26). The fact that albumin synthesis in late fetal liver is increasing at precisely the time that AFP synthesis is decreasing has led several investigators to conclude that AFP acts as a fetal serum albumin (27). However, the coding sequences of these two mRNAs do not crossreact under our hybridization conditions.

Should albumin and AFP be duplicated genes, one might expect that the basic organization of their coding sequences be maintained even though significant divergence in coding, intervening, and flanking sequences may have occurred. For example, the duplicated β major and β minor globin genes of the mouse have retained tight sequence homology in coding and a portion of the intervening sequences, whereas flanking sequences have diverged completely (28). In contrast, the α-globin gene, which diverged from the two β-globin genes much earlier, has retained no sequence homology with the β-globins. However, the organization of the gene is identical to that of the β-globin genes in that the two intervening sequences occur at the same positions (29). Thus, careful comparison of the organization of the AFP and albumin genes within a species should prove useful in assessing the likelihood that they share a common ancestry.

We thank Dimitris Kioussis, Kevin Dolan, and Paul van de Rijn for many useful and stimulating discussions in the course of this work. We are also grateful to Jon G. Seidman and Philip Leder for providing the mouse DNA library and to Margery Sullivan for her assistance in the electron microscopy studies. This work was supported by Grant CA 23572 from the National Cancer Institute, a Basil O'Connor Starter Grant, Grants CA 06927 and RR 05539 from the National Institutes of Health awarded to The Institute for Cancer Research, and an appropriation from the Commonwealth of Pennsylvania. M.B.G. is a trainee of the Medical Scientist Training Program on National Institutes of Health Grant 07170, and S.M.T. is the recipient of an American Cancer Society Junior Faculty Award.

1. Abelev, G. I. (1974) *Transplant. Rev.* **20,** 3–37.
2. Gitlin, D. & Boesman, M. (1966) *J. Clin. Invest.* **45,** 1826–1838.
3. Ruoslahti, E., Pihko, H. & Seppala, M. (1974) *Transplant. Rev.* **20,** 38–60.
4. Ruoslahti, E. & Seppala, M. (1979) *Adv. Cancer Res.* **29,** 276–346.
5. Abelev, G. I., Perova, S. D., Khramkova, N. I., Postnikova, Z. A. & Irlin, I. S. (1963) *Transplantation* **1,** 174–180.
6. Abelev, G. I., Assecritova, I. V., Kraevsky, N. A., Perova, S. D. & Perevodchikova, N. I. (1967) *Int. J. Cancer* **2,** 551–558.
7. Innis, M. A. & Miller, D. L. (1977) *J. Biol. Chem.* **252,** 8469–8475.
8. Sell, S., Thomas, K., Michalson, M., Sala-Trepat, J. & Bonner J. (1979) *Biochim. Biophys. Acta* **564,** 173–178.
9. Tilghman, S. M., Kioussis, D., Gorin, M. B., Garcia Ruiz, J. P. & Ingram, R. S. (1979) *J. Biol. Chem.* **254,** 7393–7399.
10. Polsky, F., Edgell, M. H., Seidman, J. G. & Leder, P (1978) *Anal. Biochem.* **87,** 397–410.
11. Tiemeier, D. C., Tilghman, S. M. & Leder, P. (1977) *Gene* **2,** 173–191.
12. Sternberg, N., Tiemeier, D. & Enquist, L. (1977) *Gene* **1,** 255–280.
13. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
14. Maniatis, T., Jeffrey, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 1184–1188.
15. Leder, P., Tiemeier, D. & Enquist, L. (1977) *Science* **196,** 175–177.
16. Blattner, F. R., Williams, B. G., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Furlong, L.-A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L. & Smithies, O. (1977) *Science* **196,** 161–169.
17. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, E. (1978) *Cell* **15,** 687–701.
18. Maniatis, T., Jeffrey, A. & van de Sande, H. (1975) *Biochemistry* **14,** 3787–3794.
19. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
20. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3683–3687.
21. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3961–3965.
22. Thomas, M., White, R. L. & Davis, R. W. (1976) *Proc. Natl. Acad. Sci. USA* **71,** 4579–4583.
23. Sutcliffe, J. G. (1978) *Nucleic Acids Res.* **5,** 2721–2728.
24. Sargent, T. D., Wu, J.-R., Sala-Trepat, J. M., Wallace, R. B., Reyes, A. A., & Bonner, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3256–3260.
25. Ruoslahti, E. & Terry, W. D. (1976) *Nature (London)* **260,** 804–805.
26. Ruoslahti, E. & Engvall, E. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 4641–4644.
27. Tamaoki, T., Thomas, K. & Schindler, I. (1974) *Nature (London)* **249,** 269–271.
28. Tiemeier, D. C., Tilghman, S. M., Polsky, F. I., Seidman, J. G., Leder, A., Edgell, M. H. & Leder, P. (1978) *Cell* **14,** 237–244.
29. Leder, A., Miller, H., Hamer, D., Seidman, J. G., Norman, B., Sullivan, M. & Leder, P. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 6187–6191.