THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 34, Issue of August 21, pp. 21800–21807, 1998

Printed in U.S.A.

Apolipoprotein B Gene Expression in a Series of Human Apolipoprotein B Transgenic Mice Generated with *recA*-assisted Restriction Endonuclease Cleavage-modified Bacterial Artificial Chromosomes

AN INTESTINE-SPECIFIC ENHANCER ELEMENT IS LOCATED BETWEEN 54 AND 62 KILOBASES 5′ TO THE STRUCTURAL GENE*

(Received for publication, February 24, 1998, and in revised form, June 5, 1998)

Lars B. Nielsen‡§, Debra Kahn‡, Thomas Duell¶, Heinz-Ulrich G. Weier¶, Stacy Taylor‡, and Stephen G. Young‡§‖**

From the ‡Gladstone Institute of Cardiovascular Disease, San Francisco, California 94141-9100, the ‖Department of Medicine and §Cardiovascular Research Institute, University of California, San Francisco, California 94143, and ¶Resource for Molecular Cytogenetics, Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Prior studies have established that the expression of the human apolipoprotein B (apoB) gene in the intestine is dependent on DNA sequences located a great distance from the structural gene. To identify the location of those sequences, we used *recA*-assisted restriction endonuclease (RARE) cleavage to truncate the 5′- or 3′-flanking sequences from a 145-kilobase (kb) bacterial artificial chromosome spanning the entire human apoB gene. Seven RARE cleavage– modified bacterial artificial chromosomes with different lengths of flanking sequences were used to generate transgenic mice. An analysis of those mice revealed that as little as 1.5 kb of 3′ sequences or 5 kb of 5′ sequences were sufficient to confer apoB expression in the liver. In contrast, apoB gene expression in the intestine required DNA sequences 54–62 kb 5′ to the structural gene. Those sequences retained their ability to direct apoB expression in the intestine when they were moved closer to the gene. These studies demonstrate that the intestinal expression of the apoB gene is dependent on DNA sequences located an extraordinary distance from the structural gene and that the RARE cleavage/transgenic expression strategy is a powerful approach for analyzing distant gene-regulatory sequences.

The B apolipoproteins, apoB48[1] and apoB100, play important roles in the formation of the triglyceride-rich lipoproteins (1, 2). ApoB48 is essential for the assembly of chylomicrons in the intestine, and apoB100 is essential for the formation of VLDL in the liver. When apoB gene mutations prevent the synthesis of apoB48 and apoB100, neither chylomicrons nor VLDL can be detected in the plasma, and the plasma concentrations of triglycerides and cholesterol are extremely low (3, 4).

Although the basic function of apoB (the assembly of triglyceride-rich lipoproteins) is the same in the liver and intestine, the genetic control of apoB gene expression in these two tissues is strikingly different. In transgenic mouse expression studies, an ~80-kb P1 bacteriophage (p158) spanning the entire human apoB gene (and containing 19 kb of 5′ sequences and 17.5 kb of 3′ sequences) yielded human apoB expression in the liver of transgenic mice, but expression was completely absent in the intestine (5, 6). Similarly, a P1 bacteriophage clone spanning the mouse apoB gene (and containing 33 kb of 5′-sequences and 11 kb of 3′ sequences) did not direct transgene expression in the intestine of transgenic mice (7). More recently, we identified a 145-kb bacterial artificial chromosome (BAC) clone spanning the human apoB gene and used it to generate human apoB transgenic mice. That clone, which contained 70 kb of 5′ sequences and 22 kb of 3′ sequences, conferred completely appropriate levels of apoB gene expression in the absorptive enterocytes of the intestine (8).

In the current study, we have used transgenic mouse expression studies with BACs to further define the sequences that are important for the expression of the apoB gene in the intestine. Our approach involved the use of *recA*-assisted restriction endonuclease (RARE) cleavage to delete portions of the 5′- or 3′-flanking sequences from the 145-kb BAC. The RARE cleavage-modified BACs were then used to generate multiple lines of human apoB transgenic mice. This experimental approach allowed us to document that the intestinal expression of the apoB gene is controlled by DNA sequences located an extraordinary distance upstream from the human apoB gene.

MATERIALS AND METHODS

Generating Modified BACs by RARE Cleavage—All of the BACs used in this study were derived from a 145-kb BAC, BAC(70,22), spanning the human apoB gene. BAC(70,22) contained 70 kb of 5′-flanking sequences and 22 kb of 3′-flanking sequences. To delete portions of the 5′- and 3′-flanking sequences from BAC(70,22), we deleted segments of DNA between two *Eco*RI sites or two *Hin*dIII sites. BAC(70,22) contains

* This work was supported by National Institutes of Health Grant HL47660 (to S. G. Y.); a postdoctoral fellowship from the American Heart Association, Western States Affiliate (to L. B. N.); a travel grant from the Danish Heart Association/Medical Research Council (to L. B. N.); a grant from the Director, Office of Energy Research, Office of Health and Environmental Research, U. S. Department of Energy under contract DE-AC-03-76SF00098 (to H. G. W.); and Deutsche Forschungsgemeinschaft Fellowship Du 275/1-1 (to T. D.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

** To whom correspondence should be addressed: Gladstone Institute of Cardiovascular Disease, P.O. Box 419100, San Francisco, CA 94141-9100. Tel.: 415-826-7500; Fax: 415-285-5632; E-mail: steve_young @quickmail.ucsf.edu.

[1] The abbreviations used are: apo, apolipoprotein; BAC, bacterial artificial chromosome; LCR, locus control region; kb, kilobase(s); bp, base pair(s); RARE, *recA*-assisted restriction endonuclease; PCR, polymerase chain reaction; FISH, fluorescence *in situ* hybridization; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; ATPγS, adenosine 5′-*O*-(thiotriphosphate).

more than 40 *Eco*RI and more than 40 *Hin*dIII sites. To cleave specifically at only two of the *Eco*RI sites or at only two of the *Hin*dIII sites, we used RARE cleavage (9–15). In the presence of the bacterial protein *recA*, oligonucleotides can be used to "cover up" restriction sites in a BAC clone and thereby protect them from methylation by a specific DNA methylase. After inactivation of the *recA* and the methylase, the protected restriction sites in the BAC DNA can be digested with *Hin*dIII (or *Eco*RI). RARE cleavage was performed as described previously, with some minor modifications (12, 13). Each RARE cleavage reaction contained 5× *recA* buffer (32 μl; 125 mM Tris acetate, 20 mM $MgCl_2$, 2.5 mM spermidine hydrochloride, 2 mM dithiothreitol (all from Sigma), pH 7.85), ADP (8 μl; 22 mM; Boehringer Mannheim (catalogue no. 102164)), ATPγS (8 μl; 6 mmol/liter, Boehringer Mannheim (catalogue no. 1162306)), *recA* protein (20 μl; 2 mg/ml, U.S. Biochemical Corp. (catalogue no. E70028Y/Z) or New England Biolabs (Beverly, MA) (catalogue no. 249S)), 60-mer oligonucleotide(s) (4.5 μl; 160 nmol/μl for experiments with two oligonucleotides or 320 nmol/μl for experiments with a single oligonucleotide), and distilled H_2O (to accomplish a final volume of 160 μl). These reagents were mixed and prewarmed to 37 °C for 1 min. BAC DNA (4 μg in TE (10 mmol/liter Tris-HCl, pH 7.4, 1 mM EDTA; Digene Technologies, Silver Spring, MD)) and acetylated bovine serum albumin (8 μl; 2 mg/ml; New England Biolabs) were added, and the incubation was continued at 37 °C for 20 min. *Eco*RI methylase (8 μl; 4 units/μl; New England Biolabs (catalogue no. 211S or 220L)) (or *Alu*I methylase, which methylates *Hin*dIII sites) and *S*-adenosylmethionine (8 μl; 4 units/μl; New England Biolabs) were then added. After incubating at 37 °C for 60 min, *recA* protein and methylase were denatured by incubating the mixture at 65 °C for 15 min. Methylated DNA was dialyzed on filters (VS, 0.025 μm, Millipore, Bedford, MA) against 0.5× TE for 30 min and then cleaved with *Eco*RI (or *Hin*dIII) for at least 1 h at 37 °C in the buffer provided by the supplier of the enzyme (New England Biolabs or Boehringer Mannheim).

The RARE cleavage products were separated by pulsed field electrophoresis (CHEF-mapper, Bio-Rad) on 1% low melting point agarose gels (SeaPlaque GTG-agarose, FMC Bioproducts, Rockland, ME). A gel slice (200–400 mg of gel) containing the DNA fragment was excised (14, 15) and placed in 25 ml of gelase buffer (Epicentre Technologies, Madison, WI) containing 100 mmol/liter NaCl for 30 min at room temperature. The agarose was placed in an Eppendorf tube and melted by incubation at 65 °C for 10–15 min. The tube was cooled to 42 °C, and gelase (5 units/100 μl of melted gel; Epicentre Technologies) was added. After digestion of the agarose for 1 h at 42 °C, the tube was placed on ice for 2 min and then centrifuged at 15,000 × *g* for 2 min to pellet any undigested gel. The linearized DNA (a 100-μl aliquot) was then circularized by adding 1.5 μl of T4 ligase (New England Biolabs) and 10 μl of 10× ligase buffer and incubating the mixture at 14 °C for 12–16 h. The ligation product was filter-dialyzed for 30 min against 0.5× TE; a total of 1–2 μl was used to transform 20–50 μl of electrocompetent *Escherichia coli* (strain DH10B, Life Technologies, Inc.) by electroporation. The transformations were done according to the instructions from Life Technologies, except that we used 13 instead of 18 V/cm. The transformed cells were then plated on agar plates containing chloramphenicol (12.5 μg/ml, Boehringer Mannheim) and incubated at 37 °C for 16–24 h.

In most experiments, 10 bacterial colonies were picked. To confirm that the modified BACs contained the correct inserts, the BAC clones were cleaved with *Xho*I, *Not*I, *Eco*RI, and *Hin*dIII (alone and in combination) and analyzed by conventional and pulsed field agarose gel electrophoresis. In some cases, the BACs were also analyzed by sequencing across the new *Eco*RI or *Hin*dIII restriction site generated by the RARE cleavage reaction. To prepare DNA for screening of new BAC clones, bacterial colonies were grown overnight in 4.5 ml of LB medium (Digene Diagnostics, Beltsville, MD) containing 12.5 μg/ml chloramphenicol (Boehringer Mannheim). The bacterial cultures were centrifuged for 10 min at 5500 × *g*. The bacteria were resuspended in 100 μl of buffer A (50 mM Tris, pH 8.0, 10 mM EDTA, 100 mg/ml RNase A) containing 10 μl of a lysozyme solution (10 mg/ml in buffer A, Boehringer Mannheim). After an incubation for 3 min at room temperature, the resuspended cells were lysed by adding 200 μl of 0.2 M NaOH containing 1% SDS, followed by 150 μl of a 3 M potassium acetate solution in 1.5% acetic acid. The mixture was centrifuged for 5 min at 15,000 × *g*. Finally, the DNA in the supernatant fluid was precipitated with ethanol, washed in 70% ethanol, and resuspended in 20 μl of TE. This method yielded ~4 μg of BAC DNA.

Oligonucleotides Used in RARE Cleavage Reactions—We previously cloned a P1 bacteriophage, P1–70, containing 70 kb of sequences upstream from a *Not*I site within intron 1 of the apoB gene (8); we then subcloned the *Eco*RI and *Hin*dIII fragments of the P1 clone into pBluescript SK II+ (Stratagene, La Jolla, CA). The insert from P1–70 spans from intron 1 of the human apoB gene to 70 kb 5′ of the gene (8). *Eco*RI or *Hin*dIII subclones located −5, −30, −47, −54, or −62 kb 5′ to the apoB gene were identified by quantitative DNA fiber mapping. The ends of these subclones were sequenced with T7, T3, M13 forward, or M13 reverse primers. The sequence on both sides of the restriction site of interest was obtained by sequencing back across the restriction sites using P1–70 as template. For these sequencing reactions, we used the oligonucleotide primers 1 (5′-GGA GAA GGG AGA CAG GGG GAT GG-3′ (to define the sequence surrounding the *Eco*RI site located −5 kb 5′ to the gene)), 2 (5′-TCT ATG AAA AAT CTA TGC CAA AT-3′ (to define the *Eco*RI site located −47 kb 5′ to the gene)), 3 (5′-ATG TCT GAA GGC AAG GAA GTA AT-3′ (to define the *Hin*dIII site located −54 kb 5′ to the gene)), and 4 (5′-AGC CCC AGG TAG ACC CCA TTC TC-3′ (to define the *Eco*RI site located −62 kb 5′ to the gene)). Oligonucleotide primers 5 (5′-TGT GGA ATT GTG AGC GGA TAA C-3′) and 6 (5′-GCT GGC GAA AGG GGG ATG TGC TG-3′) (corresponding to BAC vector sequences) were used to sequence across the *Hin*dIII cloning site at the 3′ and 5′ ends of BAC(70,22), respectively. All sequencing was done with the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer) according to the manufacturer's protocol, except for sequencing of BAC or P1 DNA, in which we used 2 or 4 μg of DNA in each sequencing reaction.

To clone both BAC(30,22) and BAC(54,22), we used a single 60-mer oligonucleotide that protected two *Hin*dIII sites; the first 30 bp of each oligonucleotide consisted of sequences immediately 5′ to the *Hin*dIII cloning site in BAC(70,22), while the last 30 bp consisted of sequences immediately 3′ to a *Hin*dIII site located −30 or −54 kb 5′ to the apoB gene. To generate the remainder of the BACs, we used pairs of two different 60-mer oligonucleotides, each of which spanned specific *Hin*dIII or *Eco*RI sites. BAC(54,22) was also cloned with a two-oligonucleotide approach. Oligonucleotides were purchased from Life Technologies or Oligos Etc. (Wilsonville, OR). The oligonucleotides used to generate the BACs used in this study were as follows: BAC(70,1.5), 7 (5′-ATC CTT CAA CCT CCA CCA TTC TTT CAT AAA GCT TCA CAT TCT CTT GGT GAA GCC CAC TGT-3′ (covering a *Hin*dIII site +1.5 kb 3′ of the apoB gene)) and 8 (5′-CTA TTT AGG TGA CAC TAT AGA ATA CTC AAG CTT TTG GCT CCA TTG AAG GAT TAT TTG ATG-3′ (covering the *Hin*dIII site at the 3′ cloning site within BAC(70,22))); BAC(30,22), 9 (5′-TCC TCT AGA GTC GAC CTG CAG GCA TGC AAG CTC TAT AGT TTT AAC TTT TAC ATG AAG GTC-3′ (covering both the 5′ *Hin*dIII cloning site and a *Hin*dIII site located −30 kb 5′ to the apoB gene)); BAC(54,22) (using the single oligonucleotide approach), 10 (5′-TCC TCT AGA GTC GAC CTG CAG GCA TGC AAG CTT CAT CCC CTA ACC CCA AAA AAC AAT TTA-3′ (covering both the 5′ *Hin*dIII cloning site and a *Hin*dIII site −54 kb 5′ to the apoB gene)); BAC(54,22) (using the two-oligonucleotide approach), 11 (5′-GAT CAA TGA TGA AGG AGT TGA GAA AAC AAG CTT CAT CCC CTA ACC CCA AAA AAC AAT TTA-3′ (covering a *Hin*dIII site −54 kb 5′ to the apoB gene)) and 12 (5′-TCC TCT AGA GTC GAC CTG CAG GCA TGC AAG CTT TAA ACA TGC TTT TGT GGT GTG GCT TCC-3′ (covering the 5′ *Hin*dIII cloning site in BAC(70,22))); BAC(5,22), 13 (5′-ACA TTT CCC TTC ATA GCT CTG AAC AAG GAG AAT TCA GCC CAA TTC TCA TGG CCT TCT ACA-3′ (covering an *Eco*RI site −5 kb 5′ to the apoB gene)) and 14 (5′-AGT GAA TTG TAA TAC GAC TCA CTA TAG GGC GAA TTC GAG CTC GGT ACC CGG GGA TCC TCT-3′ (covering an *Eco*RI site in the BAC vector located 49 bp from the 5′ cloning site)); BAC(62,1.5), 15 (5′-TGC TGT GTG AGC ACT GAC AGT TCA GAA TTC CTT GAA GTA TAC AGA AGG TAG GGA AGG GAA-3′ (covering an *Eco*RI site −62 kb 5′ to the apoB gene)) and oligonucleotide 13; BAC(Δ5–47,1.5), 15 (5′-ACA TAT TGA CTA TAT CCT TAA TGC TCT GAA TTC TCA TTA CAT TTG TAC AGA TGT GGG GCC-3′ (covering an *Eco*RI site −47 kb 5′ to the apoB gene)) and oligonucleotide 13; BAC(Δ5–62,1.5), oligonucleotides 13 and 15.

Generation of Transgenic Mice—BAC plasmid DNA for microinjection was prepared as described previously from 1-liter bacterial cultures that had been grown in LB medium containing 12.5 μg/ml chloramphenicol (16). To prepare the DNA for microinjection into fertilized mouse eggs, 10–30 μg of BAC DNA was cleaved with *Bss*HII (which cleaves the BAC vector twice but did not cleave within the insert of any of the BACs). The *Bss*HII-cleaved DNA was size-fractionated on a 1% low melting point agarose pulsed field gel (SeaPlaque GTG, FMC Bioproducts), and the segment of the gel containing the large *Bss*HII fragment was excised and digested with gelase (Epicentre Technologies) (17, 18). The DNA solution was then adjusted to 0.5, 2, or 3 ng/μl and microinjected into fertilized mouse eggs (C57BL/6J × SJL for the generation of two lines of BAC(5,22) mice and four lines of BAC(70,1.5) mice; FVB/N for the generation of the other transgenic lines). The

concentration of the microinjected DNA had no discernible effect on transgenesis efficiency or on transgene copy number.

Transgenic founders were identified by a human apoB-specific radioimmunoassay (5) or enzyme-linked immunosorbent assay. To determine transgene copy numbers, 5 or 10 μg of genomic DNA from human leukocytes or tails of mice was digested with *Eco*RI or *Hin*dIII. The digested DNA was separated on 1% agarose gels and blotted onto a nylon membrane. Southern blot analysis was performed with a ^{32}P-labeled 1857-bp *Bam*HI–*Eco*RI fragment from exon 26 of the human apoB gene (apoB cDNA nucleotides 4650–6507) using Quickhybe (Stratagene). Transgene copy number was assessed by comparing the intensity of the human apoB signal in transgenic mouse genomic DNA with that in human genomic DNA with a phosphor imager (BAS1000, Fujix). In some experiments, transgene copy number was determined with slot-blot analysis of genomic DNA.

PCR Analysis of the Intactness of BAC Transgenes—To determine whether any of the transgenic animals had incorporated only a portion of the BAC transgene, we used PCR to enzymatically amplify BAC vector sequences located at the far 5′ or far 3′ ends of the transgene (2). In these experiments, a 272-bp fragment of the BAC vector at the 5′ end of the transgene and a 322-bp fragment at the 3′ end of the transgene were amplified using oligonucleotide primers that have been described previously (8). PCR products were resolved on 1% agarose gels and visualized by ethidium bromide staining.

Analysis of ApoB Gene Expression by RNase Protection Assays—Total cellular RNA was isolated from mouse liver and duodenum with the Totally RNA kit (Ambion, Austin, TX). To analyze apoB expression, RNase protection assays were performed with the RPA II ribonuclease protection assay kit (Ambion) according to the manufacturer's instructions. Two human apoB antisense riboprobes were transcribed from linearized plasmids with T7 RNA polymerase (Boehringer Mannheim); one probe spanned the first 121 nucleotides of exon 1 of the apoB gene, and the other spanned 220 bp of exon 26 (cDNA nucleotides 4487–6507) (8). A 316-bp mouse glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe was used as an internal control (Ambion, catalogue no. 7431). The riboprobes were mixed with 5–25 μg of total RNA from liver and intestine. Yeast RNA (25–40 μg) was used as a negative control. After the digestion of the RNA/riboprobe mixture, the samples were electrophoresed on 6% polyacrylamide gels containing either 7 or 8.3 M urea (Novex, San Diego, CA). Dried gels were exposed to autoradiographic film to visualize protected RNA fragments.

RESULTS

Generation of RARE Cleavage–modified BACs—We recently isolated a 145-kb BAC spanning the human apoB gene (designated BAC(70,22) because it contains 70 kb of 5′-flanking sequences and 22 kb of 3′ apoB gene-flanking sequences) and showed that it conferred an appropriate pattern of apoB gene expression in both the liver and intestine of transgenic mice (8). To localize the DNA sequences that govern apoB gene expression in the intestine, we have used RARE cleavage reactions to generate a series of BAC clones containing different lengths of 5′- or 3′-flanking sequences.

We subcloned most of the *Eco*RI and *Hin*dIII fragments from the 70 kb of sequences upstream from the apoB gene and mapped their locations by quantitative DNA fiber mapping (19). After determining the DNA sequences flanking various *Eco*RI or *Hin*dIII sites, we generated oligonucleotides and used RARE cleavage reactions to cleave BAC(70,22) at specific *Hin*dIII or *Eco*RI sites within BAC(70,22). An example of a RARE cleavage reaction in which BAC(70,22) is cleaved at one *Hin*dIII site within the BAC vector and at another *Hin*dIII site 54 kb upstream from the gene is shown in Fig. 1. In that example, the 145-kb BAC was cleaved into 129- and 16-kb DNA fragments. The 129-kb DNA fragment was gel-purified, ligated, and transformed into *E. coli*, generating the modified BAC construct, BAC(54,22) (containing 54 kb 5′ and 22 kb 3′ to the apoB gene). A total of seven different RARE cleavage– modified BACs were generated and used to create human apoB transgenic mice (Fig. 2).

Transgenic Mice from BACs—The generation of human apoB transgenic mice from BAC constructs was efficient. Fifteen percent of the offspring (43 of 280 mice screened) were transgenic founders, as judged by the presence of human apoB in the plasma. The transgene copy number in the majority of founders and transgenic lines ranged from one to four (Fig. 3), although it was occasionally less than one (suggesting that a few of the founders were mosaics). Both 5′ and 3′ BAC vector sequences could be amplified from the genomic DNA of each human apoB transgenic founder animal (or F1) (Fig. 4).[2]

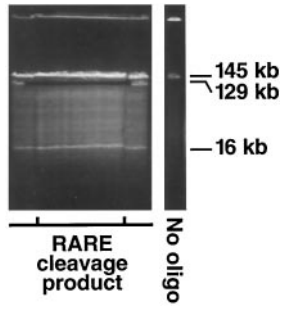

FIG. 1. **Pulsed field gel showing a RARE cleavage experiment designed to delete the sequences between −54 and −70 kb in BAC(70,22).** A single 60-mer oligonucleotide was used to protect two *Hin*dIII sites from methylation. The first 30 bp of the oligonucleotide consisted of sequences immediately 5′ to the *Hin*dIII cloning site in BAC(70, 22); the last 30 bp consisted of sequences immediately 3′ to a *Hin*dIII site located −54 kb 5′ to the apoB gene. After digestion with *Hin*dIII endonuclease, the RARE-cleaved DNA was loaded onto a wide lane of a 1% low melting point agarose gel and resolved by pulsed field gel electrophoresis. A portion of the gel was stained with ethidium bromide and aligned with the unstained portion of the gel (51). A section containing the 129-kb DNA fragment (BAC(54,22)) was removed from the unstained portion of the gel and digested with gelase. The 129-kb DNA fragment was ligated and transformed into *E. coli*. The *lane* labeled *no oligo* shows the result of an RARE cleavage reaction without the 60-mer oligonucleotide.

Localization of DNA Sequences That Control apoB Gene Expression—To determine whether the 3′-flanking sequences played a role in controlling apoB gene expression, we analyzed apoB gene expression in transgenic mice generated with BAC(70,1.5). BAC(70,1.5) differed from the parental BAC, BAC(70,22), in that it lacked the sequences 1.5–22 kb 3′ to the gene. In each of eight independent transgenic mouse lines (four F1s and four founders), the apoB gene was expressed at high levels in both the liver and intestine (Fig. 5*A*). The level of human apoB expression in the intestines of the BAC(70,15) mice appeared to be appropriate. As judged by phosphor imaging analysis, the relative human apoB expression levels in the liver and intestine paralleled those of the endogenous gene mouse apoB gene (data not shown). These data indicated that the sequences located between +1.5 and +22 kb are not important for either intestinal or liver expression of the apoB gene.

To define the 5′-flanking DNA sequences that control apoB gene expression, we initially generated mice from BACs that

[2] In prior experiments, we generated transgenic mice using BAC(120,35) and BAC(70,22) (8). In those experiments, we used two different approaches to analyze whether the BAC transgenes had been incorporated intact into the chromosomal DNA. First, we demonstrated by PCR that BAC vector sequences at either end of the transgene were present in the genomic DNA. Second, we used pulsed field gel electrophoresis of *Not*I-digested genomic DNA (obtained from hepatocyte plugs), followed by Southern blotting, to directly identify the transgene and verify that it was intact within the genomic DNA. We have also used both approaches to show the "intactness" of an 86-kb P1 bacteriophage transgene and of YAC transgenes (7, 49). Because our Southern blotting studies have invariably supported the results obtained by PCR, we performed only the PCR analyses in this study. We recognize that these PCR studies cannot provide definitive evidence for transgene intactness (50). However, in view of our prior experience, the PCR studies suggest that the BAC transgenes were incorporated intact into the genomic DNA.

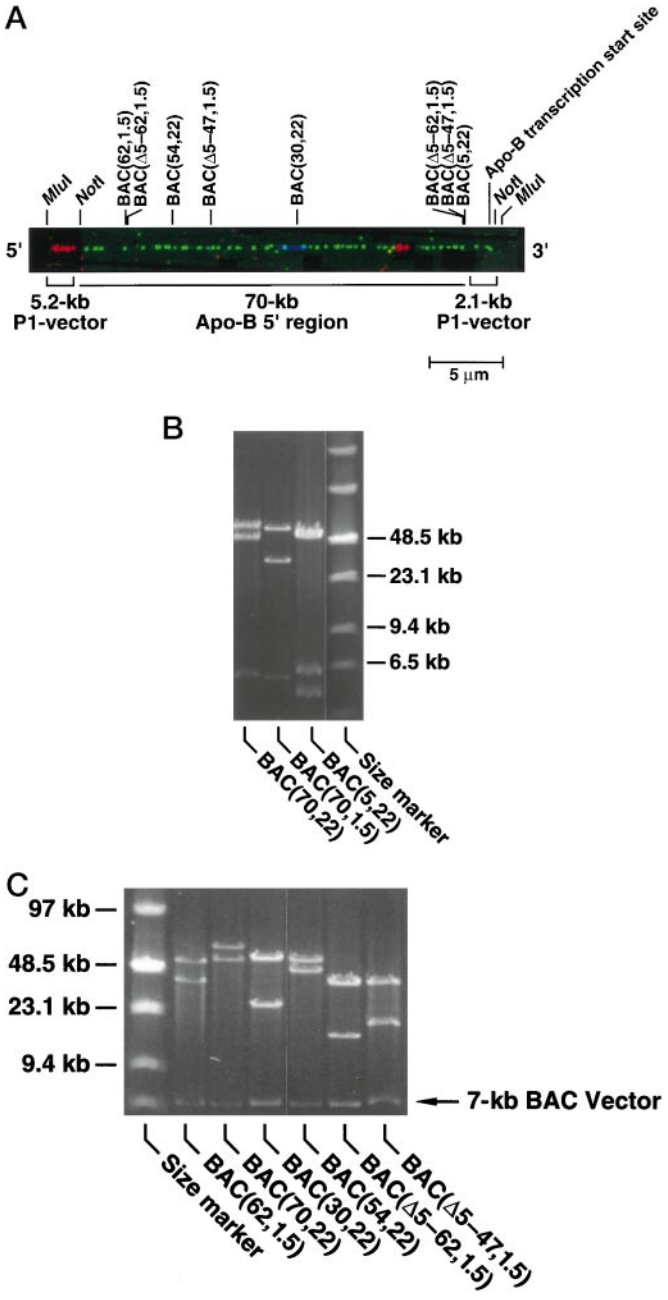

FIG. 2. *A*, map of the 5′ region of the apoB gene showing the positions of *Eco*RI or *Hin*dIII sites used in RARE cleavage reactions. The positions of these sites were mapped previously by quantitative DNA fiber mapping (19, 45). The *colored fiber* shows the results of multicolor fluorescence *in situ* hybridization (FISH) of a single DNA molecule (P1–70) that spans from intron 1 of the apoB gene to 70 kb upstream from the gene. The P1–70 DNA fiber was immobilized on glass and visualized (in *green*) by FISH, using the entire P1–70 DNA as a probe (the probe was labeled with fluorescein). Two *Hin*dIII subclones of P1–70 were labeled with digoxigenin (*blue*) and biotin (*red*), respectively. By FISH, these subclones were positioned 36 and 20 kb 5′ of the apoB gene. The orientation of the DNA fiber was revealed by FISH with a digoxigenin-labeled probe spanning a 5.1-kb portion of the P1 bacteriophage vector at the 5′ end of the DNA fiber. In some cases, the orientation of the DNA fiber was also determined with a digoxigenin-labeled probe spanning a 2.1-kb portion of the P1 bacteriophage vector at the 3′ end of the DNA fiber. *B* and *C*, ethidium bromide– stained pulsed field gels showing *Not*I digests of modified BACs generated by RARE cleavage reactions. Each of the BACs contains the entire apoB structural gene, but they differ in the lengths of 5′- and 3′-flanking sequences. The amounts of 5′- and 3′-flanking sequences in the BACs are indicated by their names (*e.g.* BAC(70,22) contains 70 kb 5′ and 22 kb 3′ to the apoB gene). Each BAC contains two *Not*I sites that flank the insert, so a *Not*I digest cuts out the 7-kb BAC vector fragment. Each BAC also contain a *Not*I site in intron 1 of the apoB gene, 221 bp 3′ to the transcription start site. Thus, digestion of the BACs results in generation of a 7-kb BAC vector fragment and two additional DNA fragments: one fragment representing the apoB 5′ region and the other representing the 43 kb of the structural gene (located downstream to the intron 1 *Not*I site) and the 3′-flanking sequences.

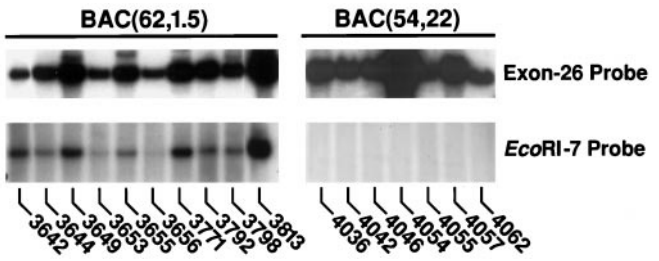

FIG. 3. **Southern blot of *Eco*RI-digested tail DNA from different lines of BAC(62,1.5) and BAC(54,22) mice.** As probes, we used a ^{32}P-labeled 1857-bp *Bam*HI–*Eco*RI fragment from exon 26 of the human apoB gene (apoB cDNA nucleotides 4650–6507) or a ^{32}P-labeled *Eco*RI fragment spanning from 58 to 62 kb 5′ to the apoB gene.

contained 5 or 30 kb of 5′-flanking sequences. Neither BAC(5,22) (F1s from two lines) nor BAC(30,22) (F1s from one line) conferred intestinal expression of the human apoB gene. Those results were consistent with data obtained with transgenic mice generated with P1 bacteriophages (*i.e.* a human apoB P1 clone containing 19 kb of 5′-flanking sequences did not yield intestinal expression (5, 6), nor did a mouse apoB P1 clone containing 33 kb of 5′-flanking sequences (7)). However, both BAC(5,22) and BAC(30,22) yielded high levels of human apoB expression in the liver, indicating that all of the sequences required for liver expression are contained within the 5 kb of sequences upstream from the structural gene (Fig. 5*B*).

To further localize the 5′ sequences controlling the intestinal expression of the apoB gene, we generated and analyzed transgenic mice with BACs containing either 54 or 62 kb of 5′-flanking sequences. Each of three BAC(62,1.5) founder mice expressed the human apoB transgene in the intestine (Fig. 6*A*), and phosphor imaging analysis revealed that the expression in the intestine relative to the liver was similar to that in the BAC(70,1.5) mice. In contrast, none of six different transgenic lines from BAC(54,22) (four founders and F1s from two lines) expressed apoB in the intestine (Fig. 6*B*). Southern blots with a probe located 58–62 kb 5′ to the gene confirmed that the BAC(62,1.5) mice contained more extensive 5′-flanking sequences than the BAC(54,22) mice (Fig. 3). We also generated transgenic mice from a BAC construct, BAC(Δ5–62,1.5), in which the DNA sequences between 5 and 62 kb upstream of the apoB gene had been deleted. These transgenic mice (one founder and F1s from one line) lacked human apoB expression in the intestine (Fig. 7). Together, these studies indicate that the sequences controlling intestinal expression of the human apoB transgene are located between 54 and 62 kb upstream from the human apoB gene.[3]

Moving the Distant Regulatory Sequences Closer to the Structural Gene by Deleting the Intervening Sequences—The sequences controlling intestinal expression of the human apoB transgene are located at a remarkable distance from the structural gene. To determine whether those distant regulatory sequences retain the capacity to confer intestinal expression when moved closer to the gene, we analyzed apoB gene expression in transgenic mice generated with a BAC in which a large

[3] Using long range PCR, we have subcloned two ~4-kb fragments, overlapping by 100 bp, that contain the DNA sequences from −54 to −62 kb 5′ of the apoB gene. These subclones are available from S. G. Y. on request.

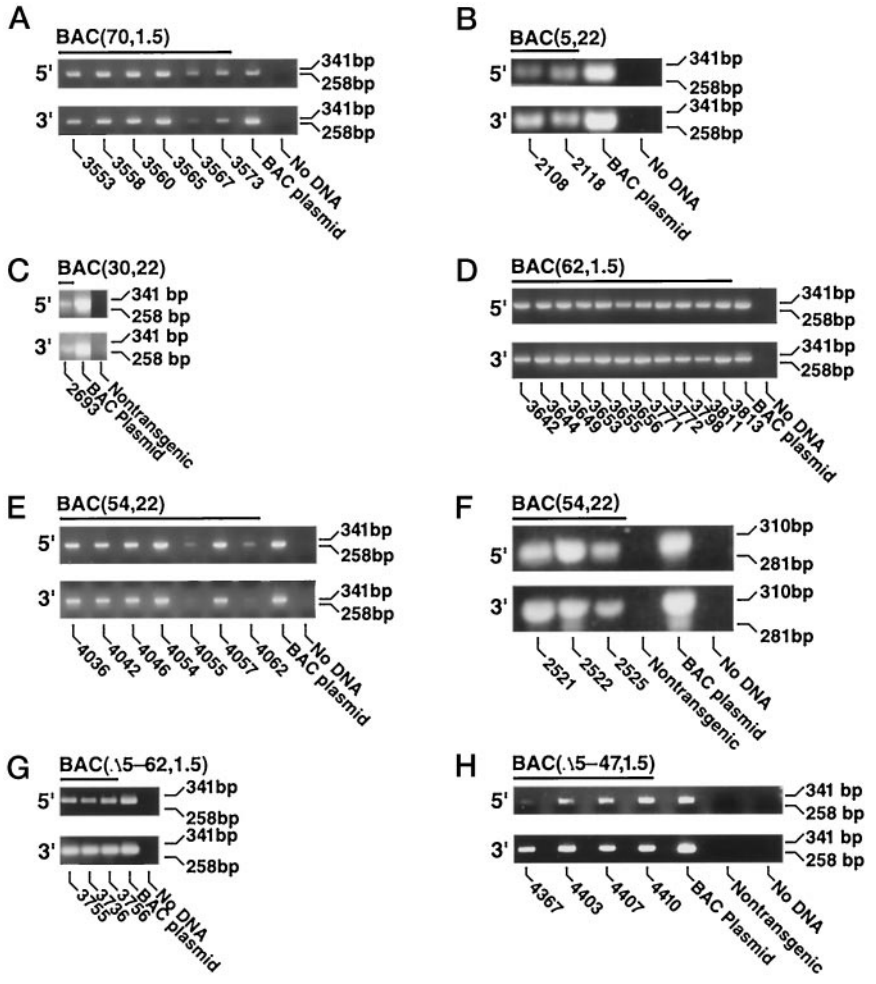

FIG. 4. **Ethidium bromide-stained agarose gels showing PCR amplification of fragments of the 3′ or 5′ ends of the BAC transgenes, using genomic DNA from transgenic founders or F1s.** Genomic DNA was isolated from tail biopsies, and BAC vector sequences located at the 5′ and 3′ ends of the transgene were enzymatically amplified using primers located within BAC vector sequences (2).

segment of the upstream sequences had been deleted. This construct, BAC(Δ5–47,1.5), was generated by deleting sequences between 5 and 47 kb upstream of the apoB gene in BAC(62,1.5). BAC(Δ5–47,1.5) yielded robust apoB gene expression in the intestine (two founders examined) (Fig. 7). These studies revealed that the sequences between −5 and −47 kb do not play an essential role in intestinal apoB gene expression and that the distant intestinal regulatory sequences function normally when moved closer to the structural gene.

DISCUSSION

Elevated plasma levels of the apoB-containing lipoproteins play a central role in the pathogenesis of atherosclerotic disease (20, 21). The relationship of apoB to atherosclerotic disease has led to intensive study of the apoB gene, including studies of the regulatory sequences that control apoB gene expression. Early attempts to define the apoB gene's regulatory sequences utilized the expression of reporter gene constructs in cultured cells (22–24). Those studies strongly suggested that the proximal promoter sequences located immediately upstream from the apoB gene were sufficient for apoB gene expression in both the liver and the intestine. Also, studies of nuclear matrix attachment sites led to the suggestion that a 47.5-kb domain from ~5.2 kb upstream to the gene to a few hundred bp 3′ to the gene contained all known *cis*-acting regulator sequences for the gene (25, 26). However, more recent transgenic animal expression studies have suggested that these studies were misleading and that the intestinal expression of the apoB gene might be controlled by distant DNA sequences (5–8, 27–29). To obtain a more detailed understanding of the regulation of apoB gene expression, particularly the DNA sequences that control apoB gene expression in the intestine, we generated apoB transgenic mice using seven different BACs containing different lengths of 5′- and 3′-flanking sequences. Analysis of the tissue-specific expression pattern of 2–8 independently generated lines from each BAC construct was unambiguous and suggested 1) that sequences located between 54 and 62 kb upstream from the apoB gene are essential for the intestinal expression of the apoB gene, 2) that these distant sequences are effective in conferring intestinal expression when they are moved closer to the structural gene by deleting intervening sequences, and 3) that the sequences located between 1.5 and 22 kb downstream from the structural gene are not important for apoB gene expression, either in the intestine or in the liver.

There are precedents for gene regulatory elements, or locus control regions (LCRs), being located at a distance from the structural gene (30–34). For example, liver expression of four genes at the apoE/CI/CIV/CI gene cluster is controlled by the same hepatic control region (15, 35–38), and expression of the ϵ-, γ-, and β-globin genes in the β-globin locus is regulated by a 5′-LCR located 5–22 kb upstream from the ϵ-globin gene (34, 39–43). Many of the previously characterized LCRs have occurred in the setting of multigene families that arose by ancient gene duplication events, and it is not difficult to imagine how these duplication events could result in a regulatory element being placed at a significant distance from the genes it controls. In contrast, the apoB gene is not known to have any adjacent family members, which makes the distance of the apoB gene's LCR (more than 54 kb from the structural gene) unique and extraordinary. It is important to note that the

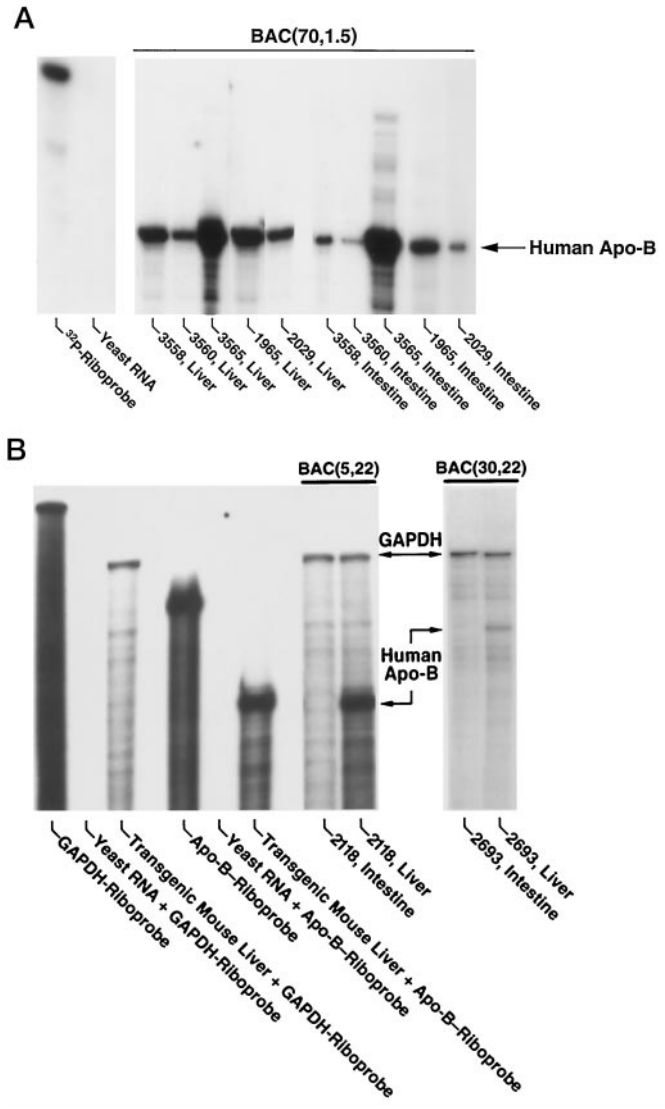

FIG. 5. *A*, RNase protection assays with a 121-bp ^{32}P-labeled human apoB riboprobe showing human apoB gene expression in liver (5 μg of RNA) and intestine (25 μg of RNA) of five lines of BAC(70,1.5) mice. *B*, RNase protection assay with 220-bp human apoB and 316-bp mouse GAPDH ^{32}P-labeled riboprobes showing human apoB and mouse GAPDH gene expression in liver (10 μg of RNA) and intestine (10 μg of RNA) of BAC(5,22) and BAC(30,22) mice. ApoB gene expression in the BAC(5,22) and in BAC(30,22) mice was analyzed in different experiments, but under similar conditions.

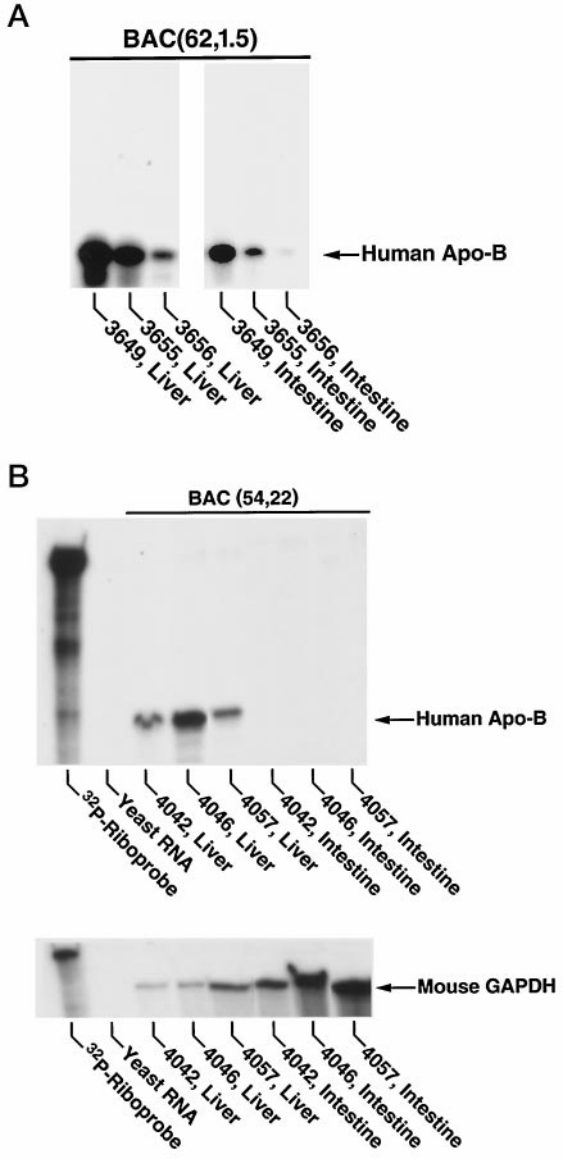

FIG. 6. *A*, RNase protection assays with a 121-bp ^{32}P-labeled human apoB riboprobe showing liver (5 μg of RNA) and intestinal (25 μg of RNA) expression of the human apoB gene in three lines of BAC(62,1.5) mice. *B*, RNase protection assay with ^{32}P-labeled 121-bp human apoB and 281-bp mouse GAPDH riboprobes showing human apoB and mouse GAPDH gene expression in the livers and intestines of three lines of BAC(54,22) mice. To analyze human apoB expression, we used 5 μg of liver RNA and 25 μg of intestinal RNA. To analyze mouse GAPDH expression, we used 2 μg of liver RNA and 2 μg of intestinal RNA.

distant regulation of apoB gene expression in the intestine is not a peculiarity of the *human* apoB gene. A mouse apoB gene clone containing 33 kb of 5′ sequences did not confer intestinal expression of the apoB gene, indicating that this distant form of apoB gene regulation has existed for more than 60 million years of mammalian evolution (44).

In our study, we used a new approach to localize distant gene regulatory sequences of large genes: mapping the gene's flanking sequences with quantitative DNA fiber mapping, using RARE cleavage to modify a large BAC clone, and then analyzing gene expression patterns in a series of BAC transgenic mice. There are several attractive features of this approach. First, high redundancy BAC libraries of human and mouse genomic DNA are commercially available. Once a clone is in hand, quantitative DNA fiber mapping (45) makes it possible to define, rapidly and accurately, useful *Eco*RI and *Hin*dIII sites for RARE cleavage reactions (19). Truncating a BAC using RARE cleavage, followed by ligating the ends of the BAC, can be performed efficiently. In addition, our studies demonstrated that the generation of transgenic mice with BAC DNA can be performed with very high efficiency, comparable with that obtained with short fragments of plasmid DNA. We believe that our approach establishes a very attractive method for delineating the regulation of large gene loci by distant *cis*-acting elements.

In the case of the β-globin gene locus, mutations in the LCRs have been shown to cause human disease. There are several examples of deletions in the 5′ LCRs that abolish β-globin synthesis and cause thalassemia, despite the fact that the protein-coding sequences of the β-globin gene remained intact (46). Therefore, we would not be surprised if mutations in the apoB gene's distant intestinal LCR cause human disease. A deletion of the apoB gene's distant LCR would be expected to eliminate chylomicron formation (leading to intestinal fat mal-

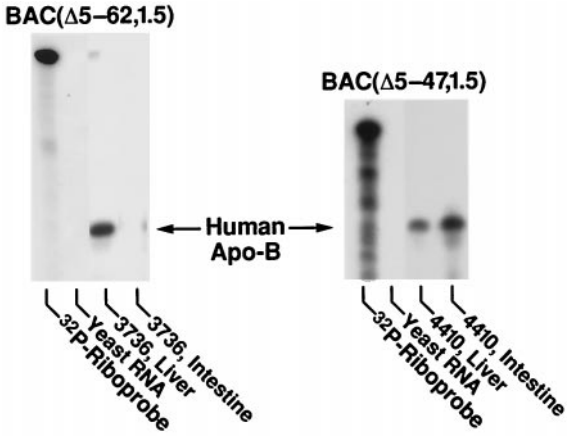

FIG. 7. **RNase protection assay with a 121-bp ^{32}P-labeled human apoB riboprobe showing liver (5 μg of RNA) and intestinal (25 μg of RNA) expression of the human apoB gene in BAC(Δ5–47,1.5) and in BAC(Δ5–62,1.5) mice.**

absorption) without affecting the ability to generate lipoproteins in the liver (28). In humans, the inability to assemble and secrete chylomicrons with a preserved capacity to secrete hepatic lipoproteins has been termed Anderson's syndrome, or chylomicron retention disease (3). In one human family, the apoB gene locus was excluded as the culprit (47). However, this syndrome is heterogeneous, both clinically and pathologically (3), and it would not be surprising if at least a few of these cases were due to mutations in the apoB gene's intestinal LCR.

While our current studies have added new insights into understanding the control of apoB gene expression, they have undeniably cast a spotlight on a series of new issues. For example, why the apoB gene's intestinal LCR is located so far away from the structural gene is not known. Nor is it known whether the distant regulatory sequences exist only to govern the apoB gene or whether they might also control the expression of another, as yet unidentified, related gene located within the upstream sequences. Finally, the precise identity of the regulatory sequences and how they act over such an enormous distance remain mysterious. Although finding answers to each of these issues will undoubtedly pose challenges, the current studies have opened doors for future investigations. First, the localization of the apoB gene's intestinal element to a fairly short stretch of DNA means that it is now feasible to embark on DNase I-hypersensitive site mapping (using tissues from transgenic mice or cultured cell lines) to more precisely identify potential regulatory sequences. Second, these studies have provided the underpinnings for a different transgenic mouse approach. Our studies demonstrated that the apoB gene's intestinal LCR can function when moved closer to the structural gene, and our mapping and subclone sequencing studies have made it possible for any laboratory to obtain the crucial regulatory sequences. These findings mean that future investigations can make use of a "co-microinjection" strategy, which is based on the fact that two DNA fragments nearly always integrate into the same site in the chromosomal DNA when they are co-microinjected into fertilized mouse eggs (8, 30, 48). For example, in the future, it will be possible to co-microinject p158 (the P1 clone containing 19 kb of 5′-flanking sequences) and a segment of DNA extending from −58 to −62 kb upstream from the apoB gene and then determine (by analyzing transgene expression patterns) whether that 4-kb segment of DNA harbors the critical sequences for intestinal apoB gene expression. The availability of new strategies for future studies on the apoB gene's regulatory sequences is welcome, since we have identified no useful sites for RARE cleavage in the sequences between −54 and −62 kb upstream from the gene.

Acknowledgments—We thank J. Gray (Mt. Zion Cancer Center, University of California, San Francisco) for pointing out the potential of quantitative DNA fiber mapping for the present studies; H. Shizuya for helping to identify the apoB-BAC clones; J. Borén for helping to define the *Hin*dIII site located 1.5 kb 3′ to the gene and for advice regarding RARE cleavage; J. Carroll and S. Gonzales for assistance with graphics; and S. Ordway and G. Howard for reviewing the manuscript.

REFERENCES

1. Havel, R. J., and Kane, J. P. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds) 7th Ed., Vol. 2, pp. 1841–1851, McGraw-Hill Inc., New York
2. Young, S. G. (1990) *Circulation* **82,** 1574–1594
3. Gregg, R. E., and Wetterau, J. R. (1994) *Curr. Opin. Lipidol.* **5,** 81–86
4. Linton, M. F., Farese, R. V., Jr., and Young, S. G. (1993) *J. Lipid Res.* **34,** 521–541
5. Linton, M. F., Farese, R. V., Jr., Chiesa, G., Grass, D. S., Chin, P., Hammer, R. E., Hobbs, H. H., and Young, S. G. (1993) *J. Clin. Invest.* **92,** 3029–3037
6. Callow, M. J., Stoltzfus, L. J., Lawn, R. M., and Rubin, E. M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 2130–2134
7. McCormick, S. P. A., Ng, J. K., Véniant, M., Borén, J., Pierotti, V., Flynn, L. M., Grass, D. S., Connolly, A., and Young, S. G. (1996) *J. Biol. Chem.* **271,** 11963–11970
8. Nielsen, L. B., McCormick, S. P. A., Pierotti, V., Tam, C., Gunn, M. D., Shizuya, H., and Young, S. G. (1997) *J. Biol. Chem.* **272,** 29752–29758
9. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graph.* **6,** 13–27
10. Ferrin, L. J., and Camerini-Otero, R. D. (1991) *Science* **254,** 1494–1497
11. Ferrin, L. J., and Camerini-Otero, R. D. (1994) *Nat. Genet.* **6,** 379–383
12. Ferrin, L. J. (1995) *Genet. Eng.* **17,** 21–30
13. Borén, J., Lee, I., Callow, M. J., Rubin, E. M., and Innerarity, T. L. (1996) *Genome Res.* **6,** 1123–1130
14. Callow, M. J., Ferrin, L. J., and Rubin, E. M. (1994) *Nucleic Acids Res.* **22,** 4348–4349
15. Allan, C. M., Taylor, S., and Taylor, J. M. (1997) *J. Biol. Chem.* **272,** 29113–29119
16. McCormick, S. P. A., Allan, C. M., Taylor, J. M., and Young, S. G. (1997) in *Transgenic Animals: Generation and Use* (Houdebine, L. M., eds) pp. 273–281, Harwood Academic Publishers, The Netherlands
17. Schedl, A., Larin, Z., Montoliu, L., Thies, E., Kelsey, G., Lehrach, H., and Schütz, G. (1993) *Nucleic Acids Res.* **21,** 4783–4787
18. Schedl, A., Montoliu, L., Kelsey, G., and Schütz, G. (1993) *Nature* **362,** 258–261
19. Duell, T., Nielsen, L. B., Jones, A., Young, S. G., and Weier, H.-U. (1997) *Cytogenet. Cell Genet.* **79,** 64–70
20. Steinberg, D., and Witztum, J. L. (1990) *J. Am. Med. Assoc.* **264,** 3047–3052
21. Young, S. G., and Parthasarathy, S. (1994) *West. J. Med.* **160,** 153–164
22. Paul, M., Wagner, J., Hoffmann, S., Urata, H., and Ganten, D. (1994) *Annu. Rev. Physiol.* **56,** 811–829
23. Paulweber, B., Brooks, A. R., Nagy, B. P., and Levy-Wilson, B. (1991) *J. Biol. Chem.* **266,** 21956–21961
24. Paulweber, B., and Levy-Wilson, B. (1991) *J. Biol. Chem.* **266,** 24161–24168
25. Levy-Wilson, B., and Fortier, C. (1989) *J. Biol. Chem.* **264,** 21196–21204
26. Levy-Wilson, B. (1995) *Prog. Nucleic Acids Res. Mol. Biol.* **50,** 161–190
27. Purcell-Huynh, D. A., Farese, R. V., Jr., Johnson, D. F., Flynn, L. M., Pierotti, V., Newland, D. L., Linton, M. F., Sanan, D. A., and Young, S. G. (1995) *J. Clin. Invest.* **95,** 2246–2257
28. Young, S. G., Cham, C. M., Pitas, R. E., Burri, B. J., Connolly, A., Flynn, L., Pappu, A. S., Wong, J. S., Hamilton, R. L., and Farese, R. V., Jr. (1995) *J. Clin. Invest.* **96,** 2932–2946
29. Fan, J., McCormick, S. P. A., Krauss, R. M., Taylor, S., Quan, R., Taylor, J. M., and Young, S. G. (1995) *Circulation* **92,** I-104 (abstr.)
30. Eccles, S., Sarner, N., Vidal, M., Cox, A., and Grosveld, F. (1990) *New Biol.* **2,** 801–811
31. Walsh, A., Azrolan, N., Wang, K., Marcigliano, A., O'Connell, A., and Breslow, J. L. (1993) *J. Lipid Res.* **34,** 617–623
32. Montoliu, L., Umland, T., and Schütz, G. (1996) *EMBO J.* **15,** 6026–6034
33. Jones, B. K., Monks, B. R., Liebhaber, S. A., and Cooke, N. E. (1995) *Mol. Cell. Biol.* **15,** 7010–7021
34. Lowrey, C. H., Bodine, D. M., and Nienhuis, A. W. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 1143–1147
35. Simonet, W. S., Bucay, N., Lauer, S. J., Pitas, R. E., Weisgraber, K. H., and Taylor, J. M. (1990) *Trans. Assoc. Am. Phys.* **103,** 119–128
36. Simonet, W. S., Bucay, N., Lauer, S. J., and Taylor, J. M. (1993) *J. Biol. Chem.* **268,** 8221–8229
37. Shachter, N. S., Zhu, Y., Walsh, A., Breslow, J. L., and Smith, J. D. (1993) *J. Lipid Res.* **34,** 1699–1707
38. Taylor, J. M., Simonet, W. S., Lauer, S. J., Zhu, G., and Walker, D. (1993) *Curr. Opin. Lipidol.* **4,** 84–89
39. Grosveld, F., van Assendelft, G. B., Greaves, D. R., and Kollias, G. (1987) *Cell* **51,** 975–985
40. Hug, B. A., Wesselschmidt, R. L., Fiering, S., Bender, M. A., Epner, E., Groudine, M., and Ley, T. J. (1996) *Mol. Cell. Biol.* **16,** 2906–2912
41. Martin, D. I. K., Fiering, S., and Groudine, M. (1996) *Curr. Opin. Genet. Dev.* **6,** 488–495
42. Caterina, J. J., Ryan, T. M., Pawlik, K. M., Palmiter, R. D., Brinster, R. L., Behringer, R. R., and Townes, T. M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 1626–1630
43. Stamatoyannopoulos, J. A., Goodwin, A., Joyce, T., and Lowrey, C. H. (1995) *EMBO J.* **14,** 106–116

44. Silver, L. M. (1995) *Mouse Genetics: Concepts and Applications*, Oxford University Press, New York
45. Weier, H.-U. G., Wang, M., Mullikin, J. C., Zhu, Y., Cheng, J.-F., Greulich, K. M., Bensimon, A., and Gray, J. W. (1995) *Hum. Mol. Genet.* **4,** 1903–1910
46. Townes, T. M., and Behringer, R. R. (1990) *Trends Genet.* **6,** 219–223
47. Strich, D., Goldstein, R., Phillips, A., Shemer, R., Goldberg, Y., Razin, A., and Freier, S. (1993) *J. Pediatr. Gastroenterol. Nutr.* **16,** 257–264
48. Hammer, R. E., Maika, S. D., Richardson, J. A., Tang, J.-P., and Taurog, J. D. (1990) *Cell* **63,** 1099–1112
49. McCormick, S. P. A., Peterson, K. R., Hammer, R. E., Clegg, C. H., and Young, S. G. (1996) *Trends Cardiovasc. Med.* **6,** 16–24
50. Young, S. G., Lusis, A. J., and Hammer, R. E. (1998) in *Molecular Basis of Heart Disease* (Chien, K. R., Breslow, J. L., Leiden, J. M., Rosenberg, R. D., and Seidman, C., eds) W. B. Saunders, Philadelphia, in press
51. McCormick, S. P. A., Linton, M. F., and Young, S. G. (1994) *Genet. Anal. Tech. Appl.* **11,** 158–164

Apolipoprotein B Gene Expression in a Series of Human Apolipoprotein B Transgenic Mice Generated with *recA*-assisted Restriction Endonuclease Cleavage-modified Bacterial Artificial Chromosomes: AN INTESTINE-SPECIFIC ENHANCER ELEMENT IS LOCATED BETWEEN 54 AND 62 KILOBASES 5′ TO THE STRUCTURAL GENE

Lars B. Nielsen, Debra Kahn, Thomas Duell, Heinz-Ulrich G. Weier, Stacy Taylor and Stephen G. Young

J. Biol. Chem. 1998, 273:21800-21807.
doi: 10.1074/jbc.273.34.21800