Vol. 166, No. 1, 1990 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1990 Pages 223-230

EVIDENCE THAT XENOPUS LAEVIS CONTAINS TWO DIFFERENT NONALLELIC INSULIN-LIKE GROWTH FACTOR-I GENES

Alan R. Shuldiner, Ajay Nirula, Laurie A. Scott and Jesse Roth

Diabetes Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 10, Room 8-S-243, Bethesda, Maryland 20892

Received November 22, 1989

Using the polymerase chain reaction (PCR), we have amplified and characterized partial nucleotide sequences of two distinct insulin-like growth factor-I genes (designated IGF-I' and IGF-I") from the amphibian, Xenopus laevis. The amplified fragments encoded much of the coding region of the mature peptide (exon III in mammalian IGF-I genes), and exhibited 93% similarity to each other, and 68-82% similarity to mammalian IGF-I amino acid sequences. Southern blot analysis using genomic DNA from a homozygous frog revealed that these two genes are nonallelic in a single organism, like the two nonallelic genes encoding Xenopus insulins that we have characterized previously. Furthermore, both IGF-I mRNAs are expressed in similar quantities in adult liver.

Insulin-like growth factor-I (IGF-I) is encoded by a member of the insulin gene family, and appears to play an important role in post-natal development (1-4). To date, the sequences of IGF-I cDNAs or peptides have been obtained only in mammals (4-9) and chicken (10), and appear to be highly conserved. Since amphibians occupy a pivotal position in evolution, bridging the present day aquatic and terrestial vertebrates (11), and since Xenopus laevis is a versatile laboratory model of vertebrate development (12), we have used the polymerase chain reaction (PCR) method (13-16) to amplify and characterize partial nucleotide sequences of two distinct Xenopus IGF-I genes (designated IGF-I' and IGF-I") which correspond to exon III in mammals. These partial sequences most likely represent two different nonallelic IGF-I genes due to tetraploidy which has been noted in this species.

MATERIALS AND METHODS

Design and synthesis of oligonucleotides. Oligonucleotides were synthesized by a Coder 300 automated DNA synthesizer (E.I. Du Pont Company; Wilmington, DE). Both upstream (sense) and downstream (antisense) PCR primers were redundant oligonucleotides derived from highly conserved regions of known mammalian IGF-I cDNAs within exon III, and spanned much of the coding region of IGF-I, including the C-, A- and D-domains, as well as a portion of the E-domain (figure 1). To facilitate subcloning of the amplified fragment, we incorporated a HindIII or EcoR1 restriction endonuclease site at the 5'-end of the upstream and downstream primers, respectively.

Polymerase chain reaction. PCR was performed using Xenopus genomic DNA (1 μg) from a homozygous frog (kindly provided by Dr. Lawrence Charnas, NIH), according to the supplier's recommendations (GeneAmp, Perkin-Elmer-Cetus Corp.; Emeryville, CA). Thirty-five cycles were

0006-291X/90 $1.50

performed; each cycle consisted of annealing (37°C, 1.5 min), extension (72°C, 1.5 min), and denaturation (94°C, 1 min) except for the last cycle in which the extension time was increased to 10 minutes to ensure completeness of extension.

Characterization of PCR amplified DNA. Twenty microliters of the reaction mixture were loaded onto a composite gel consisting of 1% agarose and 2% Nusieve GTG (GMC Bioproducts; Rockland, ME) in Tris-borate-EDTA buffer. Following electrophoresis, the gel was stained with ethidium bromide, and the products were visualized by ultraviolet transillumination. In some experiments, the gels were blotted onto Nytran filters (Schleicher and Schuell; Keene, NH), and hybridized to a radiolabeled full-length rat IGF-I cDNA probe (kindly provided by Dr. Charles T. Roberts, Jr.) (6,17). After EcoR1 and HindIII digestion (Boehringer Mannheim; Indianapolis, IN), and gel purification, the amplified band was sucloned into pGem4Z. Recombinant plasmids containing Xenopus IGF-I inserts were identified by colony hybridization, and positive colonies were isolated for sequencing using a cetyl trimethyl ammonium bromide (CTAB)(Sigma Chemical Company; St. Louis, MO) plasmid minipreparation (18). Inserts were sequenced directly from the purified plasmid using the dideoxy chain termination method (GemSeq K/RT System, Promega Biotec) according to the manufacturer's recommendations using Klenow and [^{35}S]dATP except the reactions were carried out at 50°C rather than 42°C (19). Sequences were confirmed by sequencing multiple clones from different PCR amplification reactions, as well as by sequencing each strand in both directions.

Direct sequencing of PCR fragments. One microliter of the PCR amplified product described above was reamplified asymmetrically in a second PCR reaction using identical conditions except only one primer was used (either the upstream or downstream oligonucleotide; 100 pmoles) and 25 cycles were performed (20). Excess primer was removed by ultrafiltration through a Centricon 100 ultrafiltration device, and the retentate evaporated to dryness. Denaturation was accomplished by dissolving the pellet in 0.2 M NaOH (20 μl), incubating at room temperature for 5 minutes, followed by precipitation with 5M ammonium acetate (8 μl) and cold ethanol (100 μl). (While most direct sequencing protocols do not denature the asymmetrically amplified PCR products, we have found that this step markedly improved quality and reliability). Dideoxy sequence analysis was accomplished exactly as described above using the downstream oligonucleotide as primer for the asymmetric PCR in which the upstream oligonucleotide was used, and the upstream oligonucleotide as primer for the asymmetric PCR in which the downstream oligonucleotide was used.

Genomic Southern Blot analysis. Genomic DNA (1 μg) from erythrocytes of a single homozygous Xenopus laevis was digested with either BglII, BamH1, EcoR1, or HindIII (Boehringer Mannheim). Following electrophoresis of each digest on a 1% agarose gel, Southern blot analysis (17) was performed using the 192 base pair Xenopus IGF-I' insert which was radiolabeled by random priming (21,22).

RNase Protection Assay. Total RNA was prepared from adult Xenopus liver, and from pooled Xenopus embryos from a given stage using the guanidinium thiocyanate method (23). pGem4Z containing the 192 base pair Xenopus IGF-I' insert was isolated from DH5-alpha E. coli cells using the cleared lysate method (24). Plasmid DNA was digested with HindIII, and the linearized plasmid was purified by agarose gel electrophoresis. Radiolabeled antisense Xenopus IGF-I' RNA was synthesized using SP6 polymerase (Promega Biotec) and [^{32}P]UTP according to the manufacturer's protocol. The RNase protection assay was performed essentially as described (25).

RESULTS

Genomic DNA from a homozygous frog was amplified by PCR using redundant oligonucleotides corresponding to well conserved sequences within exon III of mammalian IGF-I genes (figure 1). An amplified product of the predicted size (i.e. 192 base pairs) was obtained which hybridized strongly to a radiolabeled rat IGF-I cDNA probe (figure 2 and data not shown). (When these same oligonucloetides were used to amplify genomic DNA from rat, mouse, human, guinea pig, chicken and catfish, we obtained a very similar band with each,

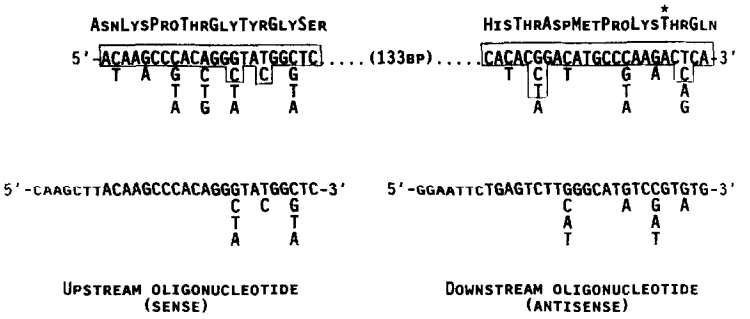

Figure 1. Design of IGF-I oligonucleotides. The upstream and downstream IGF-I oligonucleotides are 133 nucleotides apart, correspond to highly conserved regions of the IGF-I gene, and encompass much of the coding region. Depicted are all possible codons which can encode the conserved amino acid sequence. Codons that are known to occur in mammalian IGF-Is are enclosed in boxes. While it was impossible to represent all possible codons, we favored codons which were known to be used in mammals, and also purposely made the oligonucleotides redundant at their 3'-ends to increase the likelihood of an exact match at the 3'-end. The asterisk (*) indicates the nucleotide position in which porcine IGF-I contains guanine rather than adenosine resulting in an amino acid change from threonine to alanine.

which hybridized to the radiolabeled rat IGF-I cDNA probe confirming our prediction that these regions of the IGF-I gene are highly conserved across a broad range of vertebrate species; data not shown) .

The amplified DNA derived from Xenopus was digested with EcoR1 and HindIII to generate sticky ends, and subcloned into EcoR1 and HindIII digested pGem4Z. Dideoxy sequence analyses of seven independent clones revealed 192 base pair inserts that were identical in sequence to each other, and were highly similar but distinct from the known nucleotide sequences of IGF-I cDNAs or genomic clones from mouse (71% identity), pig (73% identity), rat (74% identity), human (76% identity) and chicken (77% identity) (J. Serrano, unpublished data) (figures 3 and 7).

To test whether there were two distinct nonallelic IGF-I genes in Xenopus similar to the other pairs of nonallelic genes which have been described in Xenopus (26-29), we treated DNA from a homozygous frog with four different restriction endonucleases followed by electrophoresis on agarose. Blotting and hybridization with the radiolabeled 192 base pair

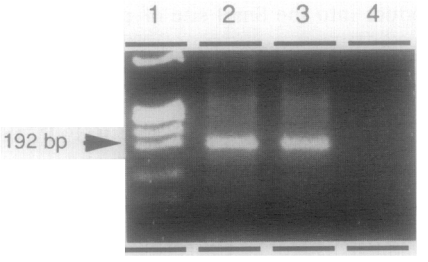

Figure 2. PCR amplification of a portion of Xenopus IGF-I genes. Lane 1, HaeIII digest of ϕX174 DNA; lane 2, amplification using genomic DNA from a homozygous frog; lane 3, amplification using genomic DNA from a heterozygous frog; and lane 4, negative control (no DNA).

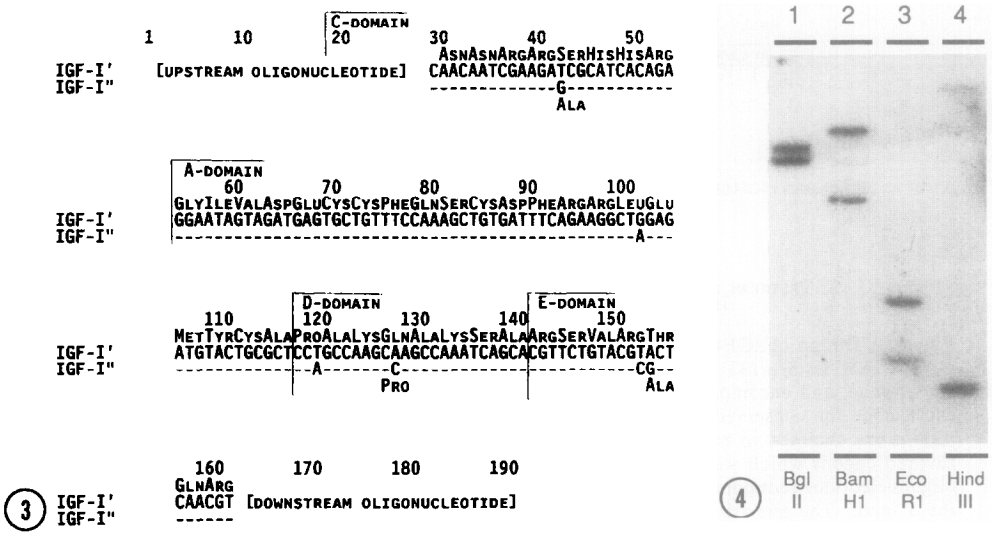

Figure 3. Partial nucleotide sequence of Xenopus IGF-I' and IGF-I" genes. The nucleotide sequence and predicted amino acid sequence of Xenopus IGF-I' are shown above, while the Xenopus IGF-I" gene is depicted below specifying the six differences in nucleotides and the three amino acid substitutions.

Figure 4. Genomic Southern blot analysis of DNA from a homozygous Xenopus. Genomic digests were hybridized to a 192 base pair radiolabeled Xenopus IGF-I' probe. See text for conditions.

Xenopus IGF-I' sequence obtained by PCR revealed two distinct bands in each digest, which we believe represents two distinct nonallelic IGF-I genes (figure 4).

Our earlier failure to detect the second IGF-I gene following subcloning raised the possibility that the redundant oligonucleotides derived as consensus sequences of known mammalian IGF-Is were selectively amplifying one Xenopus IGF-I gene, and not the other. However, when we asymmetrically amplified the 192 pase pair IGF-I fragment, and sequenced the single-stranded products directly, we obtained the same nucleotide sequence that we had obtained earlier by subcloning, except in six positions two nucleotides were observed, one of which corresponded to the original IGF-I sequence (IGF-I') while the other sequence was novel (IGF-I") (figure 3 and 5). These findings confirmed that both IGF-I genes were being amplified during PCR, and suggested that the subcloning procedure we had used favored IGF-I' over IGF-I". We therefore repeated the PCR amplification with the same oligonucleotides, but blunt-end ligated the PCR product into the SmaI site of pGem4Z. Of the nine independent clones obtained in this fashion, six were identical to the original IGF-I', while the other three were identical to the second sequence obtained by direct sequencing (IGF-I"). Of the six individual nucleotide substitutions between IGF-I' and IGF-I", three were silent and three resulted in conservative amino acid changes (figure 3).

In order to determine whether one or both Xenopus IGF-I mRNAs were expressed, we analyzed mRNA from adult frog liver with an RNase protection assay (figure 6). In this assay, we were able to detect both Xenopus IGF-I' and IGF-I" mRNAs using a single riboprobe. The undigested IGF-I' riboprobe was 202-nucleotides in length (lane 2 in figure 6) and was

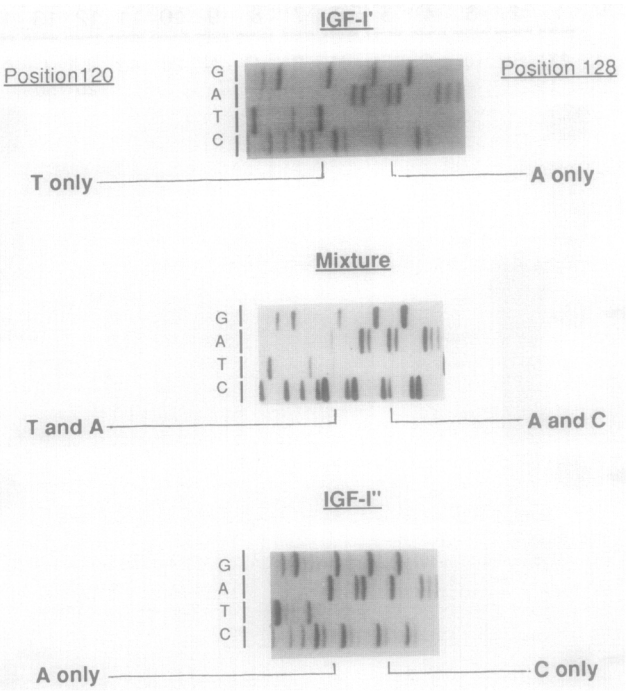

Figure 5. Simultaneous direct sequencing of two nonallelic Xenopus IGF-I genes. Depicted are two examples where direct dideoxy sequence analysis of the PCR amplified Xenopus IGF-I fragment that contained a mixture of both IGF-I' and IGF-I" DNA resulted in two nucleotides at some positions (viz. positions 120 and 128). In other experiments subcloning of the same PCR amplified product followed by dideoxy sequence analysis of independent clones allowed sequencing of each nonallelic gene separately.

completely digested in the presence of RNase (lane 3 in figure 6). Depending on how well the consensus oligonucleotides matched the true IGF-I' sequence, hybridization of the 202 nucleotide IGF-I' riboprobe to IGF-I' mRNA followed by RNase digestion would result in a protected band between 178 nucleotides in length (if the consensus oligonucleotides matched the endogenous IGF-I mRNA sequence exactly), to 133 nucleotides in length (if the consensus oligonucleotides contained mismatches at their 3'-ends). In adult liver, we observed two major protected bands, one of which was approximately 140 nucleotides in length and corresponded to Xenopus IGF-I' mRNA.

If IGF-I" mRNA were present, and RNase digestion occurred at the mismatches between the IGF-I' riboprobe and IGF-I" mRNA one would expect a partially protected band significantly smaller than 140 nucelotides. In liver, in addition to the 140 nucleotide protected band corresponding to Xenopus IGF-I' mRNA, we also observed a second band, approximately 75 nucleotides in length which corresponded to Xenopus IGF-I" mRNA. Furthermore, as determined from the relative intensities of the two bands, IGF-I' mRNA and IGF-I" mRNA were present in similar quantities in adult liver. These findings are consistent with our observations that both insulin I and II mRNAs and proteins are present in Xenopus pancreas (26,30). Using this assay, we failed to find IGF-I mRNAs in embryonic frogs, (i.e., oocytes, morula, blastula, gastrula, neurula, and 2-day, and 4-day tadpoles) even after prolonged exposure of the gels

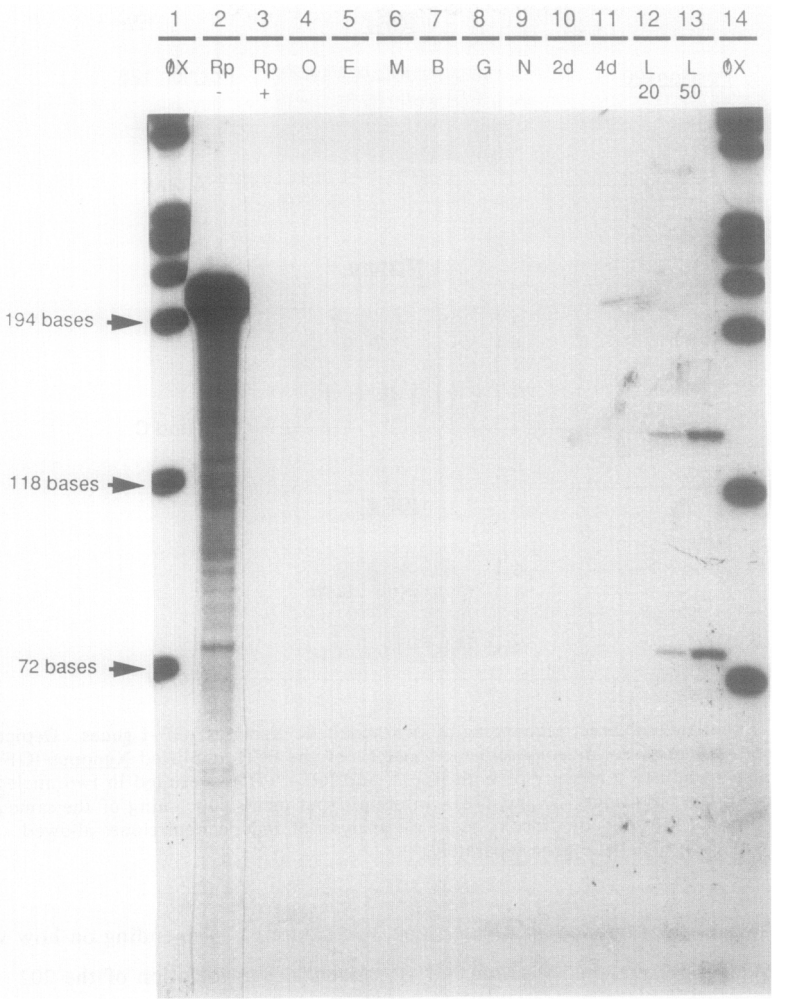

Figure 6. RNAse protection assay. Total RNA (20 to 50 μg) from unfertilized Xenopus oocytes (lane 4) and eggs (lane 5), embryos from multiple stages (lanes 6-9), and tadpoles (lanes 10 and 11) and adult liver (lanes 12 and 13) were hybridized with a 202 nucleotide antisense Xenopus IGF-I' riboprobe. Also shown is the riboprobe in the absence (lane 2) and presence (lane 3) of RNAse. A less exposed autoradiograph of the undigested probe revealed that > 98% of the radioactivity resided in the 202-nucleotide band. Lanes 1 and 14 depict a radiolabeled HaeIII digest of ϕX174 DNA.

(figure 6). These findings are consistent with patterns of IGF-I expression observed in mammals, where IGF-I appears relatively late in development (1,2,4,31,32). We plan to reexamine IGF-I expression during early embryogenesis in Xenopus using more sensitive techniques.

DISCUSSION

We have demonstrated that Xenopus laevis contains two very similar but distinct nonallelic IGF-I genes. Like the two nonallelic insulin genes in Xenopus, the two IGF-I genes are presumed to be the result of genomic duplication, which has been estimated to have occurred approximately thirty million year ago in ancestral diploid Xenopus to yield

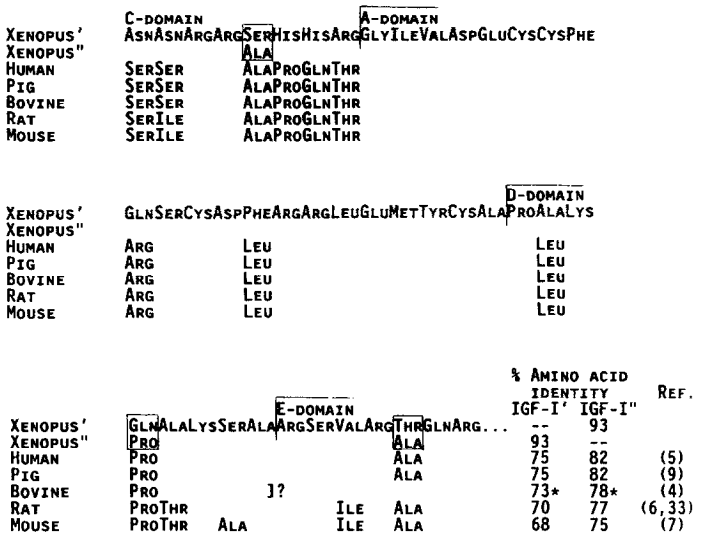

Figure 7. Comparison of Xenopus IGF-Is to IGF-Is of other species. Percent amino acid identity relative to Xenopus IGF-I' is calculated. Note the high divergence in the C-domain except for the two arginines. The asterisk (*) indicates that the calculated percent identity of bovine IGF-I to Xenopus IGF-Is does not include the E-domain since this sequence is not yet known.

present-day tetraploid species (29). Both IGF-I genes are transcribed and their respective mRNAs are present in similar quantities, at least in adult liver. Whether differential expression of these two nonallelic IGF-I genes occur in liver or extrahepatic tissues remains to be determined.

The similarity of Xenopus IGF-Is to mammalian forms of IGF-I is striking and highly conserved in the regions studied (figure 7). Interestingly, the sequence of the gene which encodes the A-domain, D-domain, and amino terminus of the E-domain showed very high conservation whereas the C-domain which is the functional equivalent of the C-peptide of insulin, showed a very high rate of substitutions, typical of insulin C-peptides, but atypical for the C-domains of the mammalian IGF-Is, which appear to be highly conserved (figure 7). That the pair of dibasic amino acids (arg-arg) in the C-domain is totally conserved between amphibian and mammalian IGF-Is is somewhat surprising, especially because this pair of dibasic amino acids is not known to be cleaved. It would appear worthwhile to test whether cleaved forms of IGF-I do occur. With the homologous probes now available, we plan to study IGF-I gene expression and its physiological role during embryogenesis in Xenopus.

ACKNOWLEDGMENTS

We would like to thank Drs. Charles T. Roberts, Jr. and Simeon I. Taylor, Mr. Neubert Philippe, and Mrs. Esther Bergman. This research project was supported in part from a grant from the Diabetes Education and Research Foundation, and the Juvenile Diabetes Foundation.

REFERENCES

1. Zapf, J., and Froesch, E.R. (1986) Hormone Res. 24, 121-130.
2. Lund, P.K., Moats-Staats, B.M., Hynes, M.A., Simmons, J.G., Jansen, M., D'Ercole, A.J., and Van Wyk, J.J. (1986) J. Biol. Chem. 261, 14539-14544.
3. Girbau, M., Gomez, J.A., Lesniak, M.A., and De Pablo, F. (1987) Endocrinology 121, 1477-1480.
4. Honegger, A., and Humbel, R.E. (1986) J. Biol. Chem. 261, 539-575.
5. Jansen, M., Van Schaik, F.M.A., Ricker, A.T., Bullock, B., Woods, D.E., Gabbay, K.H., Nussbaum, A.L., Sussenbach, J.S., and Van den Brande, J.L. (1983) Nature 306, 609-611.
6. Roberts, C.T., Jr., Lasky, S.R., Lowe, W.L., Jr., Seaman, W.T., and LeRoith, D. (1987) Mol. Endocrinol. 1, 243-248.
7. Bell, G.I., Stempien, M.M., Fong, N.M., and Rall, L.B. (1986) Nucleic Acids Res. 14, 7873-7882.
8. Shimatsu, A., and Rotwein, P. (1987) J. Biol. Chem. 262, 7894-7900.
9. Tavakkol, A., Simmen, F.A., and Simmen, R.C.M. (1988) Mol. Endocrinol 3, 674-681.
10. Dawe, S.R., Francis, G.L., McNamara, P.J., Wallace, J.C., and Ballard, F.J. (1988) Endocrinol. 117, 173-181.
11. Young, J.Z. (1981) in The Life of Vertebrates, 3rd Ed., pp.224-275, Clarendon Press, Oxford.
12. Dawid, I.B., and Sargent, T.D. (1988) Science 240, 1443-1447.
13. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) Science 230, 1350-1354.
14. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988) Science 239, 487-491.
15. Lee, C.C., Wu, X., Gibbs, R.A., Cook, R.G., Muzny, D.M., and Caskey, C.T. (1988) Science 239, 1288-1291.
16. Wilks, A.F. (1989) Proc. Nat. Acad. Sci. 86, 1603-1607.
17. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
18. Del Sal, G., Manfioletti, G., and Schneider, C. (1988) Nucleic Acids Res. 16, 9878.
19. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H., and Roe, B.A. (1980) J. Mol Biol. 143, 161-178.2
20. Innis, M.A., Kenneth, B.M., Gelfand, D.H., and Brow, M.D. (1988) Proc Nat. Acad. Sci. 85, 9436-9440.
21. Feinberg, A.P., and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
22. Lowe, W.L.,Jr., Schaffner, A.E., Roberts, C.T., Jr., and LeRoith, D. (1987) Mol. Endocrinol. 1, 181-187.
23. Chirgwin, J.M., Przbyla, A.E., MacDonald, R.J., and Rutter, W.J. (1979) Biochemistry 18, 5294-5299.
24. Guerry, P., LeBlanc, D.J., and Falkow, S. (1973) J. Bacteriol. 116, 1064-1066.
25. Lowe, W.L., Jr., Roberts, C.T., Jr., Lasky, S.R., and LeRoith, D. (1987) Proc. Natl. Acad. Sci. 84, 8946-8950.
26. Shuldiner, A.R., Phillips, S., Roberts, C.T., Jr., LeRoith, D., and Roth, J. (1989) J. Biol. Chem. 264, 9428-9432.
27. Schorpp, M., Dobbling, U., Wagner, U., and Ryffel, G.U. (1988) J. Mol. Biol. 199, 83-93.
28. Martens, G.J.M., Groenen, P.J.T.A., Braks, A.A.M., and Bussemakers, M.J.G. (1989) Nucleic Acids Res. 17, 3974.
29. Bisbee, C.A., Baker, M.A., Wilson, A.C., Hadji-Azimi, I., and Fischberg, M. (1977) Science 195, 785-787.
30. Shuldiner, A.R., Bennett, C., Robinson, E.A., and Roth, J. (1989) Endocrinology 125, 469-477.
31. Han, V.K.M., D'Ercole, A.J., and Lund, P.K. (1987) Science 236, 193-197.
32. Rotwein, P. (1986) Proc. Nat. Acad. Sci. 83, 77-81.
33. Tamura, K., Kobayashi, M., Ishii, Y., Tamura, T., Hashimoto, K., Niwa, M., and Zapf, J. (1989) J. Biol. Chem. 264, 5616-5621.