EXHIBIT 39

Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 7580–7584, November 1985
Biochemistry

Cloning and expression of the human erythropoietin gene

(erythropoietic factor/glycoprotein hormone/mixed oligonucleotide probes/genomic screening)

FU-KUEN LIN*, SIDNEY SUGGS*, CHI-HWEI LIN*, JEFFREY K. BROWNE*, RALPH SMALLING*, JOAN C. EGRIE*, KENNETH K. CHEN*, GARY M. FOX*, FRANK MARTIN*, ZIPPORA STABINSKY*, SAYED M. BADRAWI*, POR-HSIUNG LAI*, AND EUGENE GOLDWASSER†

*Amgen, 1900 Oak Terrace Lane, Thousand Oaks, CA 91320; and †Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637

Communicated by Robert T. Schimke, May 14, 1985

ABSTRACT **The human erythropoietin gene has been isolated from a genomic phage library by using mixed 20-mer and 17-mer oligonucleotide probes. The entire coding region of the gene is contained in a 5.4-kilobase *Hin*dIII–*Bam*HI fragment. The gene contains four intervening sequences (1562 base pairs) and five exons (582 base pairs). It encodes a 27-amino acid signal peptide and a 166-amino acid mature protein with a calculated M_r of 18,399. The erythropoietin gene, when introduced into Chinese hamster ovary cells, produces erythropoietin that is biologically active *in vitro* and *in vivo*.**

The interest in erythropoietin (Epo) has been documented since the turn of the century (1). Epo is the principal hormone involved in the regulation and maintenance of a physiological level of circulating erythrocyte mass (2, 3). The hormone is produced primarily by the kidney in the adult and by the liver during fetal life (4–6) and is maintained in the circulation at a concentration of ≈15–30 milliunits/ml of serum (7–9), or about 0.01 nM under normal physiological conditions. Production of Epo is stimulated under conditions of hypoxia (1, 10). Epo is proposed to exert its biological effect by attaching to specific binding sites on erythroid progenitor cells to stimulate their differentiation into mature erythrocytes (3, 11). However, when there is progressive destruction of kidney mass, such as in chronic renal failure, an anemia results due to a decreased production of Epo (12, 13). Uremic, anemic sheep and rats with decreased blood Epo levels respond to treament with Epo (14, 15). Thus, a therapeutic role for Epo appears probable in the treatment of anemia associated with renal failure.

Miyake *et al.* (16) purified Epo to homogeneity from urine of patients with severe aplastic anemia. However, it has been difficult, due to the scarcity of starting material, to obtain the amount of purified material required to investigate adequately its biological and molecular properties. We describe here the isolation of human Epo gene, based on limited amino acid sequence data from human Epo. Two probe mixtures, each containing 128-oligonucleotide sequences, were used in the library screening. This gene, when expressed in mammalian cells, encodes the production of Epo that is fully biologically active *in vitro* and *in vivo*. A preliminary report of this research has appeared (17).

MATERIALS AND METHODS

Human Genomic Library. A Charon 4A phage-borne human fetal liver genomic library prepared according to the procedure of Lawn *et al.* (18) was obtained from Tom Maniatis (Harvard University).

Construction of Oligonucleotide Probes. Purified human urinary Epo isolated from the urine of patients with aplastic anemia (16) was subjected to tryptic digestion. The resulting fragments were isolated and sequenced by using an Applied Biosystems gas-phase microsequencer (unpublished data). A hexapeptide and a heptapeptide containing the least codon degeneracy were selected for oligodeoxyribonucleotide probe synthesis. The phosphoramidite method was used for oligonucleotide synthesis (19, 20). Each probe mixture contained a pool of 128-oligonucleotide sequences. The probe mixtures were

```
Probe mixture
  EpV = Val-Asn-Phe-Tyr-Ala-Trp-Lys
     3' CAA TTG AAG ATG CGA ACC TT 5'
          T   A   A   A   T
          G               G
          C               C

Probe mixture
  EpQ = Gln-Pro-Trp-Glu-Pro-Leu
     3' GTT GGA ACC CTT GGA GA 5'
          C   T       C   T A
              G           G
              C           C
```

The probe mixtures were labeled at the 5′ end with [γ-^{32}P]ATP, 7500–8000 Ci/mmol (ICN) (1 Ci = 37 GBq), by using T4 polynucleotide kinase (21).

Hybridization Procedures. Phage plaques were amplified according to the procedures of Woo (22) except that GeneScreen*Plus* filters and NZYAM plates [NaCl, 5 g; $MgCl_2 \cdot 6H_2O$, 2 g; NZ-Amine A, 10 g; yeast extract, 5 g; Casamino acids, 2 g; maltose, 2 g; and agar, 15 g (per liter)] were utilized. Phage particles were disrupted and the DNAs were fixed on filters (50,000 plaques per 8.4 × 8.4 cm filter). The air-dried filters were baked at 80°C for 1 hr and then subjected to proteinase K digestion [50 μg of proteinase K per ml of buffer solution containing 0.1 M Tris·HCl (pH 8.0), 0.15 M NaCl, 10 mM EDTA, and 0.2% $NaDodSO_4$] for 30 min at 55°C. Prehybridization with a 1 M NaCl/1% $NaDodSO_4$ solution was carried out at 55°C for 4 hr or longer.

The hybridization buffer contained 0.025 pmol/ml of each of the 128 probe sequences in 0.9 M NaCl/5 mM EDTA/50 mM sodium phosphate, pH 6.5/0.5% $NaDodSO_4$/100 μg of yeast tRNA per ml. Hybridization was carried out at 48°C for 20 hr by using the EpV probe mixture. This is 2°C below the lowest calculated dissociation temperature (t_d) (23) for members of the mixture. At the completion of hybridization, the filters were washed three times with 0.9 M NaCl/90 mM sodium citrate, pH 7.0/0.1% $NaDodSO_4$ at room temperature

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: Epo, erythropoietin; CHO, Chinese hamster ovary; DHFR, dihydrofolate reductase; kb, kilobase(s); bp, base pair(s); nt, nucleotide(s); SV40, simian virus 40; t_d, dissociation temperature.

and two or three times with 0.9 M NaCl/90 mM sodium citrate, pH 7.0/1% $NaDodSO_4$ at the hybridization temperature, 10 min per wash. Prior to hybridization with the second probe mixture, the filters were incubated at 100°C in 0.15 M NaCl/15 mM sodium citrate, pH 7.0/0.1% $NaDodSO_4$ for 2 min to remove the hybridized probes. The filters were again prehybridized as described above and then hybridized with the EpQ mixed probes at 46°C (4°C below the lowest calculated t_d for this mixture) and washed as described above.

Assembly of Expression Vector for the Epo Gene. For direct expression of the genomic Epo gene, the 4.8-kilobase (kb) *Bst*EII–*Bam*HI fragment of λHE1 (see *Results*), which contains the entire Epo gene, was used. After converting the *Bst*EII site into a *Bam*HI site with a synthetic linker, the fragment was inserted into the unique *Bam*HI site of the expression vector pDSVL (unpublished data), which contains a dihydrofolate reductase (DHFR) minigene from pMg1 (24). The resulting plasmid pDSVL-gHuEPO (Fig. 1*A*) was then used to transfect Chinese hamster ovary (CHO) $DHFR^-$ cells (25) by the calcium phosphate microprecipitate method (26). The transformants were selected by growth in medium lacking hypoxanthine and thymidine. The culture medium used was Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, penicillin, streptomycin, and glutamine (25).

Epo Assays. *In vitro* and *in vivo* biological activities were determined by using cultured rat bone marrow cells (27) and exhypoxic polycythemic mice (28), respectively. The radioimmunoassay for Epo used antibody raised against 1% pure human urinary Epo and a human urinary Epo preparation of 1100 units/mg (CAT-1) as standard.

Isolation of Epo mRNA. For obtaining Epo mRNA, the 5.4-kb *Hin*dIII–*Bam*HI restriction fragment from λHE1 (see *Results*) was inserted into a shuttle vector, pSV4ST (unpublished data). The resulting chimeric plasmid pSVgHuEPO (Fig. 1*B*) was used to transfect COS-1 cells (ATCC CRL1650) by the calcium phosphate microprecipitate method (26). After culture for 3 days, RNA was prepared from the transfected cells by the guanidinium thiocyanate procedure of Chirgwin *et al.* (29) and poly(A)$^+$ mRNA was isolated by binding to oligo(dT)-cellulose (30).

cDNA Cloning. An Epo cDNA bank was constructed according to a modification of the general procedures of Okayama and Berg (31) by using the poly(A)$^+$ mRNA described above (unpublished data).

Southern Blotting. Human lymphocyte DNA was digested to completion by various restriction enzymes. The digested DNA samples were electrophoresed on a 0.7% agarose gel and transferred to a GeneScreen*Plus* filter by a modification of the Southern procedure (32): after the gel was denatured with 0.5 M NaOH/1.5 M NaCl, it was rinsed briefly with distilled water and then transferred with 1.5 M NaCl/0.15 M sodium citrate, pH 7.0. The filter was probed with a nick-translated ^{32}P-labeled human Epo cDNA clone that contained the coding region from the *Bst*EII site to the poly(A) tail region (see Fig. 3).

DNA Sequencing. Restriction fragments were cloned into M13 phage vectors by using *Escherichia coli* strains JM103 or JM109 as host (33) and were sequenced by the dideoxy method of Sanger *et al.* (34). Some regions were sequenced by kinase labeling or end-fill labeling of restriction fragments followed by chemical cleavage as described by Maxam and Gilbert (21).

RESULTS

Isolation and Characterization of the Epo Gene from a Human Genomic Library. A human fetal liver genomic library in bacteriophage vector Charon 4A was screened for the gene coding for Epo. Two pools of mixed synthetic oligonucleotides were used as probes as described in *Materials and Methods*. A library of 1,500,000 phage clones was screened sequentially with both probe mixtures. The 20-mer probe mixture hybridized to 272 phage plaques and the 17-mers hybridized to ≈4000 plaques. Only 4 plaques hybridized with both probe mixtures. Subsequent Southern blot and DNA sequence analyses confirmed that three of the four clones that hybridized with both probe mixtures contain at least a portion of the Epo gene. One clone, λHE1, which contained the complete Epo gene, was chosen for further analysis.

The restriction endonuclease map of the human Epo gene from clone λHE1 is shown in Fig. 2. The protein coding region of the gene is divided by four intervening sequences. Since the transcription initiation site of the mRNA for Epo has not been determined due to lack of human tissue mRNA,

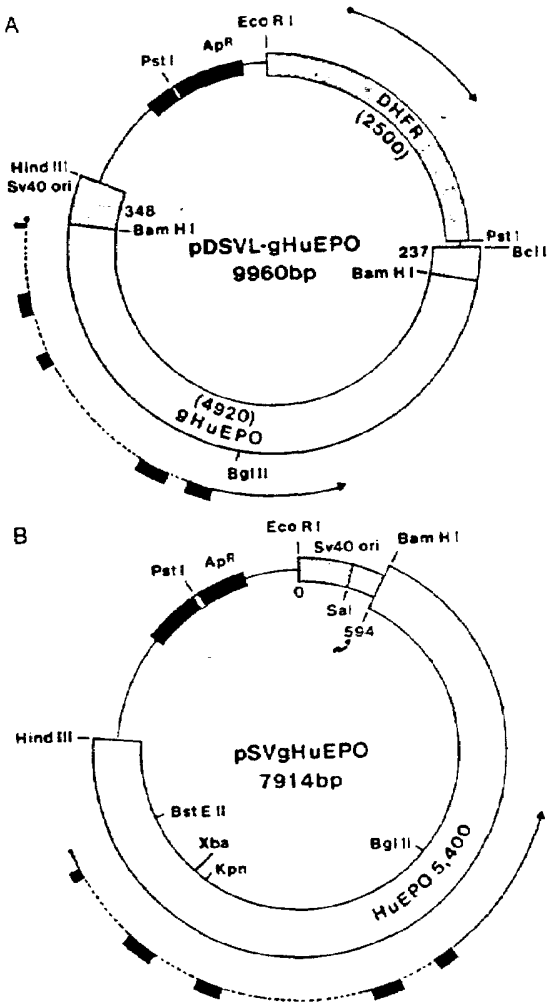

FIG. 1. Assembly of Epo expression plasmids. (*A*) pDSVL-gHuEPO contains a DHFR minigene as a *Eco*RI–*Pst* I fragment from pMg1 (obtained from R. Schimke), simian virus 40 (SV40) origin of replication and early/late promoters in the *Hin*dIII–*Bam*HI fragment, SV40 nucleotides (nt) 2538–2770 in the *Bam*HI–*Bcl* I fragment, pBR322 nt 2448–4362 (obtained from R. Tjian) in the *Hin*dIII–*Eco*RI fragment, and the *Bst*EII–*Bam*HI fragment of the Epo gene. The SV40 late promoter is used to drive the expression of the Epo gene. (*B*) pSVgHuEPO contains the SV40 origin of replication, early/late promoters, and early poly(A) signals in the *Eco*RI–*Bam*HI fragment, pBR322 nt 2448–4362 in the *Hin*dIII–*Eco*RI fragment, and the *Hin*dIII–*Bam*HI fragment of the Epo gene. Arrows indicate the orientation of transcription. bp, Base pairs.

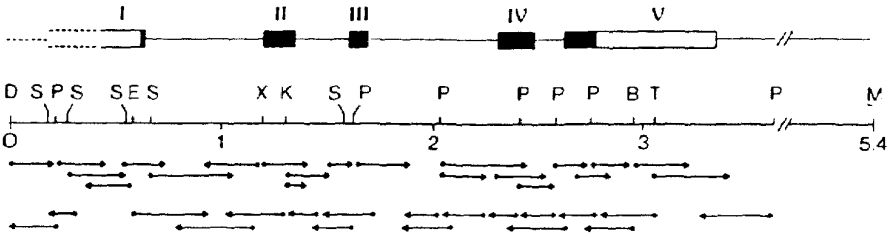

FIG. 2. Restriction map of the human Epo gene. Exons I–V are indicated by boxes. The solid boxes denote the regions of the exons that are translated. The dashed lines on the 5' side of exon I indicate that the start site for transcription is unknown. The restriction endonuclease recognition sites are abbreviated as follows: B, *Bgl* II; D, *Hin*dIII; E, *Bst*EII; K, *Kpn* I; M, *Bam*HI; P, *Pst* I; S, *Sma* I; T, *Sst* I; and X, *Xba* I. Map distances are shown in kb. Arrows below the map indicate the regions that were sequenced.

the boundary on the 5' side of exon I is undefined. Restriction sites that were utilized in the sequence analysis are also shown in Fig. 2 as are the segments that were sequenced.

Fig. 3 shows the nucleotide sequence of the human Epo gene. The exons were identified by comparison of the nucleotide sequence to the amino acid sequence of human urinary Epo (unpublished data) and by comparison of the genomic sequence to the cDNA sequence derived from mRNA isolated from CHO cells producing recombinant Epo. The exon–intron boundaries of the Epo gene conform to consensus splice rules (35).

The Epo gene encodes a protein of 193 amino acids. Based on the NH_2-terminal amino acid sequence of purified Epo, the last 166 residues correspond to the mature protein with a calculated M_r of 18,399 in an unglycosylated form. The sequence of the first 27 amino acids, predominantly hydrophobic residues, is consistent with this region encoding a leader peptide (36). The amino acid sequence starting at position +1 corresponds to the sequence of the amino terminus of expressed recombinant Epo product in CHO cells (data not shown). As indicated in Fig. 3, the mature protein has three potential sites for N-linked glycosylation, one each in the second, third, and fourth exons of the gene, according to the rule of Asn-Xaa-Ser/Thr (37).

A search of the entire 626-bp region upstream from the protein initiation codon ATG did not reveal any promoter-like sequences, such as an ATA box, CCAAT box, or −100 region (38), nor were any such sequences found elsewhere in the entire gene.

There is a 565-bp untranslated region at the 3' end of the last exon. The nucleotide sequence upstream from the poly(A) site in the Epo gene does not contain the consensus poly(A) signal sequence AATAAA or any related sequences normally found at this location (39–41). Similarly, a consensus poly(A) signal is not found in the Epo cDNA clone from cynomolgus monkey (unpublished data). The only sequence resembling AATAAA is AAGAAC, found 11–13 nucleotides upstream from the poly(A) site (Fig. 3).

In the intervening sequence between exons III and IV is a member of the *Alu* family of repeated sequences. This region is 70% homologous to the consensus *Alu* sequence (42). As is typical of these sequences, this region is flanked by an imperfect direct repeat.

Expression of the Epo Gene. Biologically active recombinant human Epo was produced in CHO cells, which had been stably transformed with an expression vector containing the genomic Epo gene insert driven by the SV40 late promoter (Fig. 1*A*). A representative sample of 5.5-day conditioned medium from pDSVL-gHuEPO-transfected cells contained 18.2 units of Epo per ml when measured by the radioimmunoassay and 15.8 ± 4.6 and 16.8 ± 3.0 units/ml when measured by the *in vitro* and *in vivo* assays, respectively. The close agreement between the results of these three assays demonstrates that the Epo produced by recombinant techniques is fully biologically active. The secreted Epo has an apparent M_r of 34,000 when analyzed in an electrophoretic transfer blot. Endo-β-*N*-acetylglucosaminidase F treatment reduced the M_r of recombinant Epo from 34,000 to ≈19,000 (unpublished data), indicating that the protein is glycosylated. The recombinant Epo has been determined to contain sialic acid by gas chromatography (unpublished data). The terminal sialic acid of Epo carbohydrate structure has been shown to be required for *in vivo* activity (43).

Genomic Organization of the Epo Gene. Restriction fragment analysis of human lymphocyte DNA, probed with human Epo cDNA, revealed a single band in digests of *Bam*HI, *Eco*RI/*Hin*dIII, *Eco*RI/*Bam*HI, and *Hin*dIII/*Bam*HI. There are three visible bands in a *Pst* I digest (Fig. 4). The size of the hybridized bands of *Hin*dIII/*Bam*HI and *Pst* I are similar to those in the isolated Epo genomic clone λHE1, as shown in Fig. 2. Hybridization at lower stringency (55°C) did not reveal any additional bands, indicating that there are no other closely related Epo genes or pseudogenes. A computer-aided homology search of the human and monkey Epo genes against GenBank and the entire Dayhoff protein bank did not reveal significant homology with any published DNA or protein sequences, including angiotensinogen, which has been suggested as a possible Epo precursor (44).

DISCUSSION

Mixtures of short synthetic oligonucleotide probes have been utilized to isolate specific clones from cDNA libraries (45, 46); however, mixed oligonucleotide probes have never been used previously for the isolation of genes from mammalian genomic libraries, principally due to the complexity of the genome. Utilizing two mixtures of 128 sequences, 17 and 20 nucleotides long, we have rapidly isolated the human Epo gene from a genomic library using a "two-site" confirmation approach. This approach for screening a genomic library eliminates a great number of false positives associated with the single-probe mixture approach. In the present study, three of four clones that hybridized with both probe mixtures were authentic Epo clones, as confirmed by DNA sequence or Southern blot analyses. The high accuracy of this technique represents a great saving in the time and effort required to isolate a gene of interest.

The key to the success of the screening approach was in the optimization of various steps in the hybridization. Use of the GeneScreen*Plus* filter resulted in more efficient binding of DNA than obtained with nitrocellulose filters and also has the advantage of repeated use. A rich medium such as NZYAM was required to support good amplification of phages on this type of filter; omission of Casamino acids or maltose resulted in weaker hybridization signal. The proteinase K digestion step greatly reduced the nonspecific background, which made probing with a mixture of 128 sequences possible. Under the chosen stringent hybridization condition (2–4°C

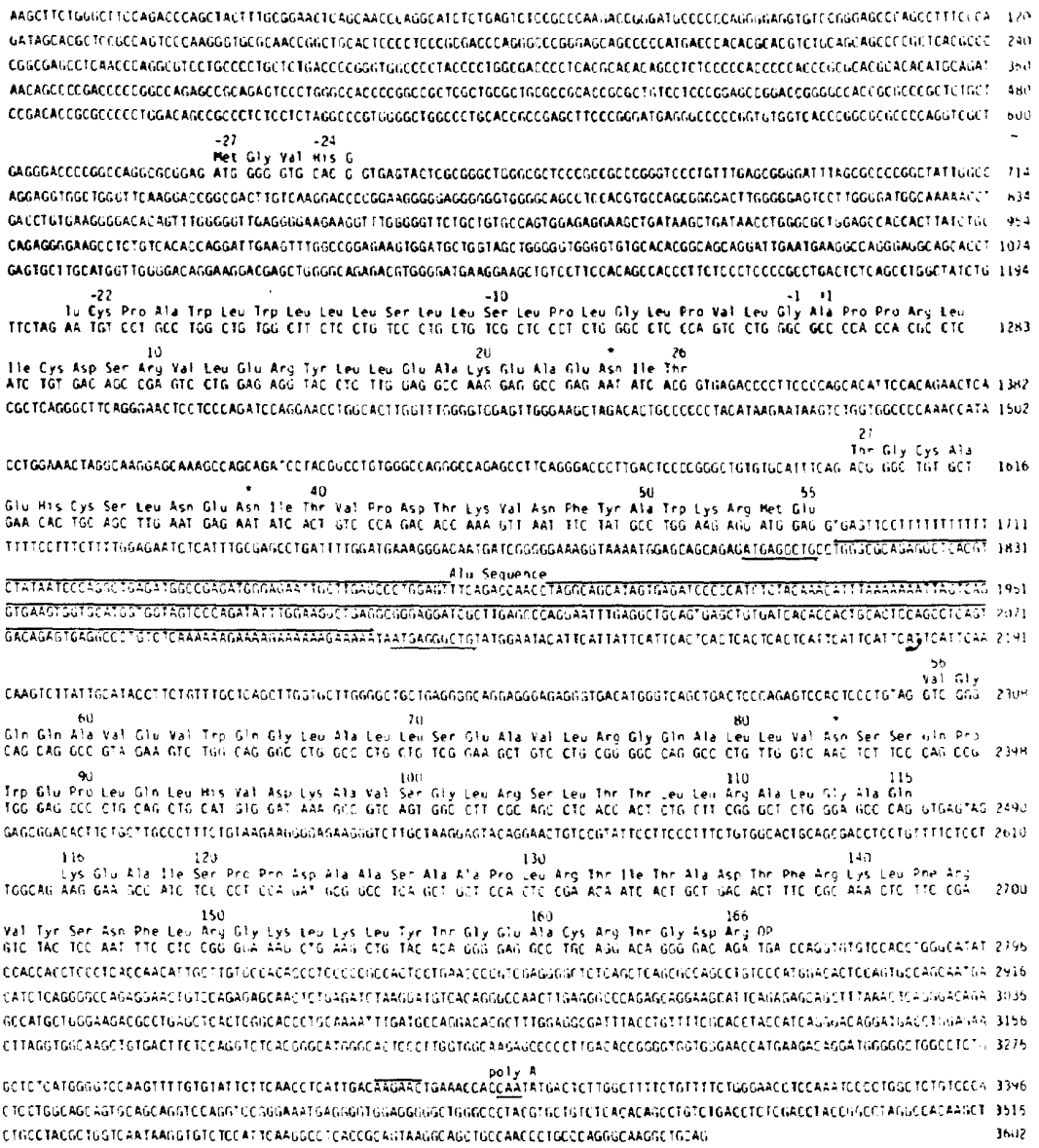

FIG. 3. Nucleotide sequence of the human Epo gene. The coding regions of the exons have been translated and the encoded amino acid sequence is shown above the nucleotide sequence. The amino acid residues are numbered and the sites of potential N-linked glycosylation are designated by asterisks. Arrows above the nucleotide sequence indicate the region that is related to the *Alu* family of repeated sequences. The arrows below the nucleotide sequence denote the direct repeat flanking the *Alu* sequence. The site that is polyadenylylated in the mRNA is underlined. Poly(A) could be added after either one of the three possible positions. The potential poly(A) signal AAGAAC is overlined.

below t_d), probe concentrations of 0.025 pmol/ml of each sequence approach the optimum required to achieve a good signal-to-noise ratio when using mixtures of this size. These optimization steps together made the genomic screening with short oligonucleotides possible.

The feasibility and ease of using large mixtures of short oligonucleotide probes to isolate mammalian genomic genes opens a new avenue to the isolation of a gene for which no known mRNA or antibody source is available and for which only limited amino acid sequence is known. Subsequently, the gene can be expressed in a mammalian expression vector, which, in turn, allows for the isolation of the mRNA and construction of cDNA libraries. The availability of the cDNA sequence provides an accurate assignment of protein coding region and exon–intron splice junctions.

The proof that the isolated genomic clone contains the Epo gene is based on the results obtained when this DNA was inserted into a SV40 promoter-containing plasmid. This plasmid, when transfected into CHO cells, led to the production of a protein having the biological activity of Epo—i.e., the protein stimulates erythrocyte production. This is evident from *in vivo* bioassay results based on the enhanced ^{59}Fe incorporation into heme in erythrocytes of mice treated with the protein and on an elevated hematocrit value when the protein was injected into normal mice (unpublished data).

It has been proposed that biologically inactive forms of Epo, termed erythropoietinogen and proerythropoietin, are produced in the kidney (see ref. 3 for review). Analysis of the DNA sequence does not support such a hypothesis. The Epo gene encodes a preprotein probably comprised of a 27-amino acid signal peptide and a 166-amino acid mature protein. The mature protein has been shown to be biologically active.

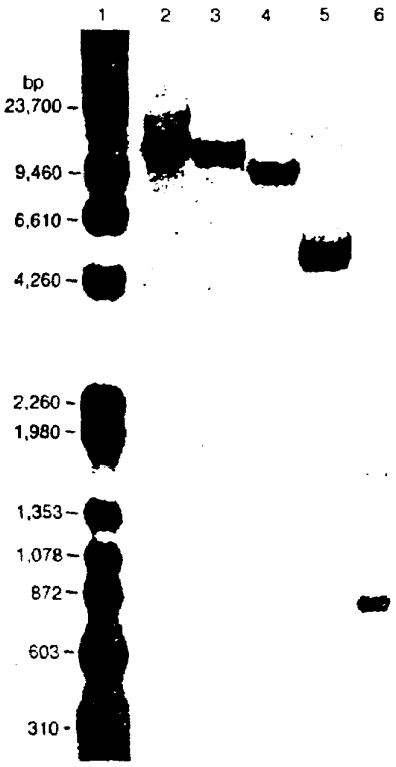

FIG. 4. Southern hybridization of human lymphocyte DNA digested with restriction endonuclease. Lane 1, ^{32}P-labeled fragments of λ*Hin*dIII digest and φX174 *Hae* III digest as molecular weight markers; lane 2, *Bam*HI; lane 3, *Hin*dIII + *Eco*RI; lane 4, *Bam*HI + *Eco*RI; lane 5, *Bam*HI + *Hin*dIII; lane 6, *Pst* I. The digested DNA was subjected to electrophoresis on a 0.7% agarose gel, transferred to a GeneScreen*Plus* filter, and hybridized with nick-translated ^{32}P-labeled Epo cDNA at 65°C. The posthybridization wash was done at 65°C with 0.15 M NaCl/15 mM sodium citrate, pH 7.0/1% $NaDodSO_4$.

There is no evidence from the gene sequence that any other form of processing is required.

The availability of sufficient quantities of recombinant Epo will facilitate a more complete understanding of the structure and function of this molecule in hemopoiesis and investigation of its potential use as a therapeutic agent for anemia patients.

After the completion of this manuscript, Jacobs *et al.* (47) reported the isolation of a human Epo gene, using a mixed oligonucleotide probe approach similar to ours, which encodes an Epo protein with amino acid sequence identical to ours. Previously, Lee-Huang (48) reported the isolation of a human Epo cDNA clone from a human cDNA bank using monoclonal antibody. Since no sequence information was given in her paper, we cannot compare this clone with ours.

We thank Dr. Tom Maniatis for the human genomic DNA library, B. Bacheller, Mary Carter, M. Castro, R. Everett, Ming Hu, Sylvia Hu, Donna Langley, T. Jones, D. Murdock, and A. Thomason for technical assistance, and Dr. Peter Dukes for the *in vivo* bioassays. We are indebted to Jennie Caruthers, Joan Bennett, Jeanne Fitzgerald, and Pat Korecky for preparation of the manuscript, and to D. Vapnek and N. Stebbing for critical reading of the manuscript.

A part of the work in E.G.'s laboratory was supported by a grant to E.G. from the National Heart, Lung and Blood Institute (HL21676).

1. Carnot, P. & Deflandre, C. (1906) *C. R. Hebd. Seances Acad. Sci. Ser. M* **3**, 384–387.
2. Goldwasser, E. (1975) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **34**, 2285–2292.
3. Graber, S. E. & Krantz, S. B. (1978) *Annu. Rev. Med.* **29**, 51–66.
4. Jacobson, L. O., Goldwasser, E., Fried, W. & Plzak, L. (1957) *Trans. Assoc. Am. Physicians* **70**, 305–317.
5. Fried, W. (1972) *Blood* **40**, 671–677.
6. Zanjani, E. D., Ascensao, J. L., McGlave, P. B., Banisadre, M. & Ash, R. C. (1981) *J. Clin. Invest.* **67**, 1183–1188.
7. Koeffler, H. P. & Goldwasser, E. (1981) *Ann. Intern. Med.* **97**, 44–47.
8. Garcia, J. F., Ebbe, S. N., Hollander, L., Cutting, H. O., Miller, M. E. & Cronkite, E. P. (1982) *J. Lab. Clin. Med.* **99**, 624–635.
9. Cotes, P. M. (1982) *Br. J. Haematol.* **50**, 427–438.
10. Erslev, A. J. (1955) *Blood* **10**, 954–961.
11. Krantz, S. B. & Goldwasser, E. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7574–7578.
12. Brown, R. (1965) *Br. Med. J.* **2**, 1036–1038.
13. Erslev, A. J., Caro, J., Miller, O. & Silver, R. (1980) *Ann. Clin. Lab. Sci.* **10**, 250–257.
14. Anagnostou, A., Barone, J., Kedo, A. & Fried, W. (1977) *Br. J. Haematol.* **37**, 85–91.
15. Eschbach, J. W., Mladenovic, J., Garcia, J. F., Wahl, P. W. & Adamson, J. W. (1984) *J. Clin. Invest.* **74**, 434–441.
16. Miyake, T., Kung, C. K.-H. & Goldwasser, E. (1977) *J. Biol. Chem.* **252**, 5558–5564.
17. Lin, F. K., Lai, P. H., Smalling, R., Egrie, J., Browne, J., Lin, C. H., Wang, F. F. & Goldwasser, E. (1984) *Exp. Hematol.* **12**, 357 (abstr.).
18. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
19. Beaucage, S. L. & Caruthers, M. H. (1981) *Tetrahedron Lett.* **22**, 1859–1862.
20. McBride, L. J. & Caruthers, M. H. (1983) *Tetrahedron Lett.* **24**, 245–248.
21. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
22. Woo, S. L. C. (1979) *Methods Enzymol.* **68**, 389–395.
23. Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K. & Wallace, R. B. (1981) in *Developmental Biology Using Purified Genes*, eds. Brown, D. D. & Fox, C. F. (Academic, New York), pp. 683–693.
24. Gasser, C. S., Simonsen, C. C., Schilling, J. W. & Schimke, R. T. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6522–6526.
25. Urlaub, G. & Chasin, L. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4216–4220.
26. Wigler, M., Silverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y.-C. & Axel, R. (1977) *Cell* **11**, 223–232.
27. Goldwasser, E., Eliason, J. F. & Sikkema, D. (1975) *Endocrinology* **97**, 315–325.
28. Cotes, P. M. & Bangham, D. R. (1961) *Nature (London)* **191**, 1065–1067.
29. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
30. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
31. Okayama, H. & Berg, P. (1982) *Mol. Cell. Biol.* **2**, 161–170.
32. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
33. Messing, J. (1983) *Methods Enzymol.* **101**, 20–78.
34. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
35. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
36. Watson, M. E. E. (1984) *Nucleic Acids Res.* **12**, 5145–5164.
37. Marshall, R. D. (1974) *Biochem. Soc. Symp.* **40**, 17–26.
38. Dierks, P., Van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J. & Weissman, C. (1983) *Cell* **32**, 695–706.
39. Donehower, L., Huang, A. & Hager, G. (1981) *J. Virol.* **37**, 226–238
40. Nevins, J. R. (1983) *Annu. Rev. Biochem.* **52**, 441–466.
41. Simonsen, C. C. & Levinson, A. D. (1983) *Mol. Cell. Biol.* **3**, 2250–2258.
42. Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedmann, T. & Schmid, C. W. (1981) *J. Mol. Biol.* **151**, 17–33.
43. Lowy, P. H., Keighley, G. & Borsook, H. A. (1960) *Nature (London)* **185**, 102–103.
44. Fyhrquist, F., Rosenlof, K., Gronhagen-Riska, C., Hortling, L. & Tikkanen, I. (1984) *Nature (London)* **308**, 649–652.
45. Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. & Itakura, K. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6613–6617.
46. Singer-Sam, J., Simmer, R. L., Keith, D. H., Shively, L., Teplitz, M., Itakura, K., Gartler, S. M. & Riggs, A. D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 802–806.
47. Jacobs, K., Shoemaker, C., Rudersdorf, R., Neill, S. D., Kaufman, R. J., Mufson, A., Seehra, J., Jones, S. S., Hewick, R., Fritch, E. F., Kawakita, M., Shimizu, T. & Miyake, T. (1985) *Nature (London)* **313**, 806–810.
48. Lee-Huang, S. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2708–2712.