Volume 16 Number 23 1988 Nucleic Acids Research

Identification of two novel members of *erbA* superfamily by molecular cloning: the gene products of the two are highly related to each other

Nobuyuki Miyajima, Yasunori Kadowaki, Shin-ichi Fukushige[1], Sho-ichi Shimizu, Kentaro Semba, Yuji Yamanashi, Ken-ichi Matsubara[1], Kumao Toyoshima and Tadashi Yamamoto

Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minatoku, Tokyo 108 and [1]Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565, Japan

Received September 15, 1988; Revised and Accepted November 8, 1988 Accession nos X12794, X12795

ABSTRACT

Two v-erbA-related genes, named ear-2 and ear-3, have been identified in the human genome and characterized by cDNA cloning. These genes are predicted to encode proteins that are very similar in primary structure to receptors for steroid hormones or thyroid hormone (T_3). In addition, amino acid sequences of the ear-2 and ear-3 gene products are very similar each other especially at the DNA binding domain (86% homology) and at the putative ligand binding domain (76% homology). Northern hybridization with ear DNA probes of RNAs from various tissues of a human fetus reveals that the expression of ear-2 is high in the liver whereas the expression of ear-3 is relatively ubiquitous. Hybridization analysis of DNAs from sorted chromosomes shows that the ear-2 gene is located on chromosome 19 and ear-3 on chromosome 5, indicating that the two genes are clearly different from each other.

INTRODUCTION

The avian erythroblastosis virus (AEV) induces erythroblastosis in susceptible chickens and transforms erythroblasts in vitro. In addition, AEV occasionally causes sarcomas in vivo and transforms fibroblasts in vitro. The ES4 and R strains of AEV carry two host cell-derived genes, termed v-erbA and v-erbB (1,2). The transforming potential of AEV has been shown to reside on the v-erbB gene (3,4), and v-erbA does not by itself induce tumors, but enhances the erythroblast-transforming ability of v-erbB (4).

The v-erbB gene is derived from the chicken EGF (epidermal growth factor) receptor gene (5,6) and the v-erbA gene from the cellular gene encoding the receptor for the thyroid hormone T_3 (7). Thyroid hormone-receptor complexes in the nucleus influence the transcription of various genes, e.g. the growth hormone gene (8). These facts suggest that expression at the cell

are described in the text and were labeled with ^{32}P by nick translation to a specific activity of $2X10^8$ cpm/μg of DNA. Phage DNAs were prepared as described (9) and the inserts were recloned into plasmid vector pUC19 at the EcoRI site.

Nucleotide sequence analysis

The nucleotide sequence was determined by the dideoxy chain termination method (22) with modifications (23,24) in conjunction with plasmid pUC18, pUC19 or bacteriophage M13mp18.

Blot-hybridization analysis of DNA and RNA

High molecular weight DNAs were digested with restriction endonucleases under the conditions recommended by the supplier (Takara Shuzo, Kyoto) and fractionated by electrophoresis in a 1% agarose gel. The fragments were subjected to Southern blot hybridization at 42°C for 16hr as described (9). RNAs were prepared by the guanidium isothiocyanate/cesium chloride method (25) from tissues (brain, lung, liver, kidney) of a 12-week-old single human fetus, a human placenta at term, TIG-1 cells and A431 cells. Polyadenylated (poly(A)$^+$) RNAs selected by oligo(dT)-cellulose column chromatography were denatured with 50% formamide/2.2M formaldehyde, and 5 μg of each RNA sample was subjected to electrophoresis in 1% agarose gel containing 2.2M formaldehyde (26). RNAs on the gel were subjected to blot hybridization under stringent conditions, as described (27). The DNA probes used for hybridization were labeled with ^{32}P by nick translation to a specific activity of $2X10^8$cpm/μg DNA.

Chromosome mapping

The chromosomal location of the *ear* genes was determined by Southern blot analysis of sorted chromosomes of human cell lines or from human-mouse hybrid cells, as described (28). GM0131 and GM0089 cells were provided by the Human Genetic Mutant Cell Repository (Institute for Medical Research, Camden, N.Y.).

RESULTS

Cloning and characterization of cDNA of human erbA homologs.

RNA blot hybridization analysis with the ^{32}P-labeled v-*erbA* probe (probe I, Fig. 1c), which was prepared from the cloned AEV-R genomic DNA (29), at lower stringency revealed two species

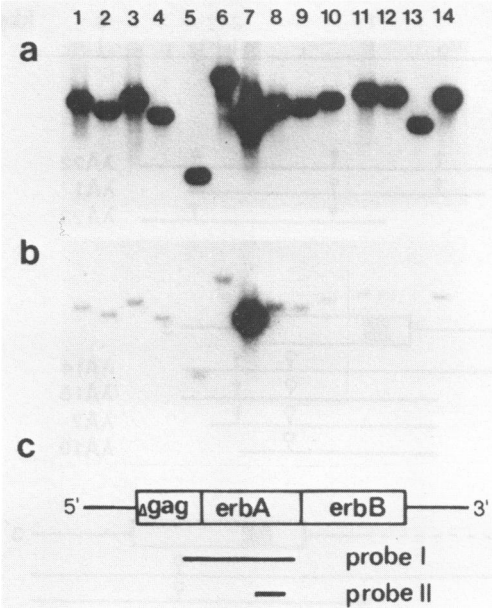

Figure 1. Southern blot analysis of cDNA inserts.
DNAs (3μg) of the recombinant phages that were isolated by screening a cDNA library of human embryo fibroblasts, TIG-1, in λgt10 were digested with EcoRI and the digests were subjected to Southern blot hybridization using probe I (a) or probe II (b) specific for the v-erbA sequence. Only 14 samples (lanes 1 to 14) are shown in this Figure. The cDNA clones in lane 1, lanes 2 and 14, lanes 3, 6 and 11, and lane 7 belongs to the ear1, ear2, ear3, and ear7 genes. Other cDNA clones in lanes 4,5,8,9,10,12 and 13 are not well characterized. Probe I was the 1.4-kbp AvaI-SstI fragment containing most of the v-erbA sequence and 400-kb of the gag gene, and probe II the 0.5-kbp PstI fragment generated from the cloned AEV-R DNA (c). Probe I, but not probe II, contained a sequence encoding the putative DNA binding domain of the v-erbA protein.

of c-erbA mRNA (2.7-kb and 5.0-kb) in human embryo fibroblasts (data not shown).

A cDNA library was constructed in λgt10 vector using mRNAs from human-embryo fibroblasts. On screening 3.5 X 10^5 independent clones of this library with the probe I under conditions of low stringency, 21 positive clones were obtained. The DNAs of these clones were digested with EcoRI and subjected to Southern hybridization analysis (Fig. 1a; only 14 clones are

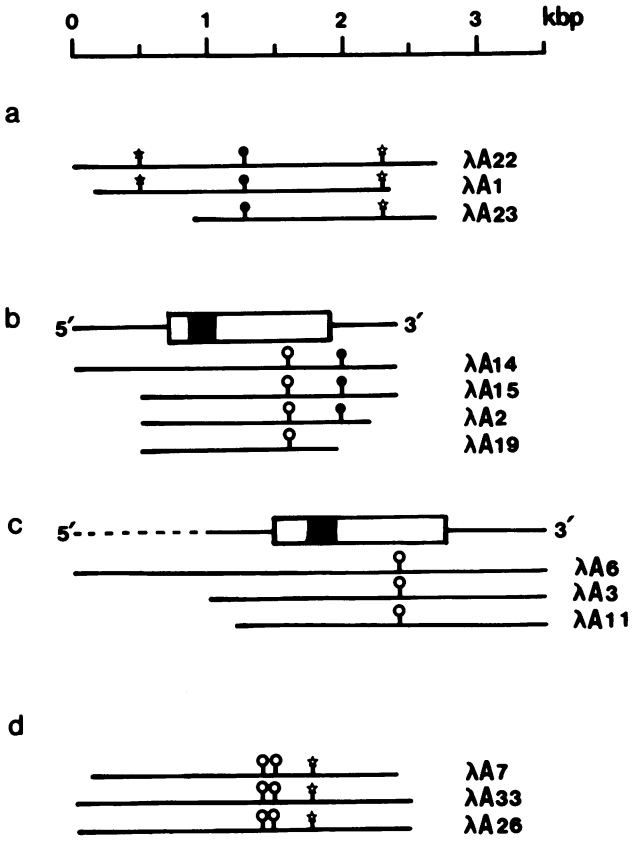

Figure 2. Schematic illustration of the cDNA clones of the ear genes.
DNA inserts and their restriction maps of cDNA clones for the ear-1 (a), ear-2 (b), ear-3 (c) and ear-7 genes are illustrated. The coding regions and non-coding regions predicted from the nucleotide sequences of the ear-2 and ear-3 cDNA clones (see Fig. 4) are shown by boxes and lines, respectively. The dotted line in (c) represents the region that was not sequenced. The closed boxes indicate the sequences encoding the putative DNA binding domain. Restriction mapping was performed by hybridizing the DNAs (3μg) of recombinant phages, which were digested with endonucleases BamHI (★), XhoI (●), EcoRI (☆) and HincII (o), with probe I (Fig. 1c). The EcoRI sites at the two ends of the inserts are not shown. The restriction maps match those predicted from the nucleotide sequence (see Fig. 4 for ear-2 and ear-3 and unpublished data for ear-1 and ear-7).

shown here). The EcoRI inserts of various length of the cDNA clones all hybridized with probe I to similar extents, except the 2.5-kbp insert of λA7 which hybridized slightly stronger than the inserts of the other clones. With the 500-bp PstI fragment prepared from AEV-R DNA (probe II in Fig. 1c), the DNA insert of λA7 hybridized strongly and those of the other clones hybridized weakly (Fig. 1b).

Restriction mapping and mutual hybridization of the DNA inserts, which had been subcloned into the plasmid vector pUC19 at the EcoRI site (data not shown), revealed that the 21 cDNA clones represented at least four erbA-related genes (Fig. 2). These were named ear-1, ear-2, ear-3 and ear-7 (ear: erbA related gene). The above hybridization data showed that the ear-7 gene represented by the λA7 insert was most closely related to the v-erbA gene. Of the 21 clones, four were of ear-2, seven clones of ear-3, one of ear-1 and one of ear-7. The other clones, which have not yet been well characterized may represent other erbA-related genes. To obtain additional cDNA clones of ear-1 and ear-7, we screened the same library with the ^{32}P-labeled cDNA inserts of λA1 and λA7 and obtained 13 positive clones. Restriction mapping analysis in conjunction with Southern blot hybridization of the inserts with the screening probes showed that two clones (λA22, λA23) were of ear-1 and two (λA26, λA33) of ear-7 (see Fig. 2).

Partial nucleotide sequencing of the cDNA clones revealed that the four erbA-related genes coded for proteins that contained sequences of 66 to 68 amino acid residues characteristic of the DNA binding domain of the chicken v-erbA/T_3-receptor gene. This sequence is also conserved in receptors for steroid hormones such as progesterone, estrogen and glucocorticoid, as well as in the T_3 receptor (Fig. 3). The high homology observed between chicken c-erbA and ear-7 (94%) suggested that the ear-7 gene was the human counterpart of the chicken c-erbA gene and corresponded to the human erbA1 gene reported previously (30). The marked homology between the DNA binding domains of ear-2 and ear-3 (85%) is also noteworthy, which prompted us to first characterize the ear-2 and ear-3 genes.

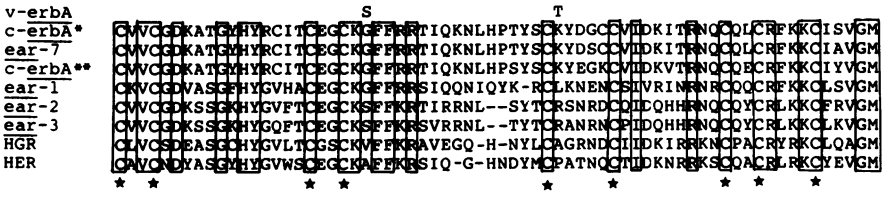

Figure 3. Alignment of the amino acid sequences of the DNA binding domain.
Amino acid sequences of the DNA binding domain of the ear-1, ear-2, ear-3, and ear-7 proteins were compared with those of the corresponding portions of the v-erbA, chicken c-erbA and human c-erbA gene products as well as of the estrogen receptor and glucocorticoid receptor. Only non-identical amino acids compared with chicken c-erbA are shown for v-erbA. The residues conserved in these proteins are boxed. Cysteine residues conserved are indicated at the bottom. c-erbA*, chicken c-erbA; c-erbA**, human c-erbAβ.

Sequence analysis of ear-2 and ear-3.

The 2380-nucleotide sequence of ear-2 was determined by sequencing the λA14 and λA2 inserts. An open reading frame of 1209-bp starting with the ATG codon at position 703 was identified (Fig. 4a). The flanking nucleotides of this ATG codon match the Kozak consensus sequence (31) and 249-bp preceding the ATG codon is an in-frame stop codon TAG. The 66 amino acid residues that are encoded by the nucleotide sequence from the TGC codon at position 868 to the ATG codon at position 1065 constitute a putative DNA binding domain homologous to that of v-erbA. The 5' untranslated 704 bp sequence of the ear-2 mRNA is suggested to form an extremely stable secondary structure with a calculated free energy (32) of -390.0 kcal because of its high content of GC. The postulated structure is shown in Fig. 4d. The typical consensus sequence AATAAA for the polyadenylation signal, but not the poly(dA) was found at extreme 3' terminus of the λA14 insert. The primary translation product of the ear-2 gene was calculated to have a molecular weight of 42,900.

Next, the 2511-nucleotide sequence of the λA3 insert of the ear-3 gene was determined. An open reading frame of 1251-bp starting with the first ATG codon at position 478 was identified (Fig. 4b). The flanking nucleotides of this ATG codon match the

surface of the v-erbB protein, a truncated form of the EGF receptor, and transcriptional deregulation caused by v-erbA, an altered form of the T_3 receptor, co-operate in complete transformation of erythroblasts.

A v-erbB-related gene other than the EGF receptor gene was recently identified in the human genome and termed c-erbB-2 (9). Nucleotide sequencing of c-erbB-2 cDNA clones suggested that this gene encoded a putative receptor protein similar to the EGF receptor (10,11,12). The ligand of this protein has yet to be identified. Similarly, several erbA-related genes have been identified in the human genome. The human counterpart of the v-erbA gene has been mapped at chromosome 17q11-21 (13,14). Other erbA homologs were mapped at 17q21.3 and 17q25 (15,16). These data indicate that there are multiple erbA-related genes on human chromosome 17. Besides these genes, at least three erbA homologs were located on human chromosome 3; one of them was shown to encode a thyroid hormone receptor (17) and one a retinoic acid receptor (18,19). The function and nucleotide sequences of the other c-erbA homolog(s) are unknown. To establish their function and structure, we screened a human cDNA library for erbA homologs. In this paper, we report the molecular cloning and characterization of novel erbA homologs of human origin.

MATERIALS AND METHODS

Isolation of cDNA clones

A cDNA library from human embryo fibroblasts, TIG-1 (20), was constructed with phage λgt10 as a vector, as described (21). The library was screened with ^{32}P-labeled DNA probe in 4X SSC (1X SSC: 0.15M NaCl/15mM sodium citrate), 50mM Hepes (pH 7.0), 20ug/ml of denatured salmon sperm DNA, 10X Denhardt solution (1X Denhardt: 0.02% polyvinylpyrrolidone/0.02% ficoll/0.02% bovine serum albumin) and 30% (relaxed conditions) or 50% (stringent conditions) formamide at 42°C for 16hr. After hybridization, the filters were washed with 2X SSC, 0.1% SDS at room temperature and then with 0.6X SSC (relaxed) or 0.1X SSC (stringent) and 0.1% SDS at 50°C. Plaques with positive signal were purified by successive plaque hybridization. The probes used

Kozak consensus sequence and 204-bp preceding the ATG codon is an in-frame stop codon. The nucleotide sequence from the TGC codon at position 733 to the ATG codon at position 928 encodes the putative DNA binding domain. It is uncertain whether the AATAAA sequence present near the extreme 3' terminus of the λA3 insert is used as poly(A) addition signal. If it is, the 5' noncoding sequence must be extremely long (about 2.5-kbp), since the ear-3 mRNA are 4.6-kb and 4.8-kb long. The free energy of a possible secondary structure at the 5' untranslated region (32) was calculated to be -207.4 kcal. Thus, this region could form a stable structure (Fig. 4d), which suggests important role for this sequence in translation of the ear-3 mRNA. The primary translation product of the ear-3 gene was predicted to have a relative molecular weight of 46,155.

The predicted amino acid sequence of not only the DNA binding domains of the ear-2 protein and ear-3 protein but also the rest of the sequence, especially the sequences which include putative ligand-binding domains, show marked homology between the two (Fig. 5). The homology of the amino acid sequences of the putative ligand-binding domain encoded by the nucleotide sequence from position 1174 to 1911 of ear-2 and that from position 1003 to 1746 of ear-3 is 76%.

Chromosomal locations of the ear-2 and ear-3 genes.

Metaphase chromosomes were prepared from two human cell lines, GM0131 and GM0089, whose karyotypes and chromosomal abnormalities have been defined. The chromosomes were separated into six or eight fractions using a fluorescence-activated cell sorter. DNA from each fraction was prepared and subjected to Southern blot hybridization after digestion with restriction endonucleases (Fig. 6). The ear-2 specific probe, a 2.4-kbp EcoRI insert from λA14, gave a positive signal in a fraction G of karyotypically normal lymphoblast GM0131 cells (Fig. 6A-a). The fraction G contained chromosomes 19 and 20. To determine which of these chromosomes carries the ear-2 gene, chromosomal DNAs from the GM0089 cell line, which carries reciprocal translocation t(X;19), were subjected to Southern blot hybridization. The ear-2 probe gave two positive signals (Fig. 6B-a), one in fraction "d", in which two translocated chromosomes, (Xpter >

Xq22 :: 19q133 > 19qter) and (19pter > 19q133 :: Xq22 > Xqter) were recovered, and the other in a fraction "e", in which chromosome 19 was recovered. Therefore, the ear-2 gene was allocated to chromosome 19. This conclusion was consistent with the finding that the ear-2 probe did not hybridize with DNAs of

a

```
   1 CTGTGTGTCAAACCGGCCGCAGCGCGAGGCCGGCGCGCGGGGTCGGTGGCGGTGGCGGCGGCGAGAGCGAGGCTTGGGCCCAGGCGCAGGCCCAGGCCAGGCAGGAGCGGCCGCAGCGGG
 121 GGCGGCAGCTCCAGAGGCGCCGGCCCAGCGGCCCTGCAGGCCCGGAGCGGCAGCCGCGGGCGCCGGGGGACCCAGGGGGCGCGGCGATCACCCCGGCGCGCGCCGCCGCCCTCCCCCC
 241 CGCCACGGGCGGCCGCGCAACTTGCGGCCAAACTTTCCCCCCGGCCTCCCGCGCTCGGATGGCGGCCCGCTGAGTGGCCGCAGCCCCCGGGCGAGCGGCGGCCAGGCCAGAGGACGCCCC
 361 CCTGCGCGGCTCGGGCCGGGCCATGGCCGCGCGCCCCCGCCGGGGCCCCAGGGCGGCGGGACGGAGTCCATCTGGGCCCCCGGCCCGCGCCTCTAGCGTGCCCCTCCCGGCCCCGGGCCC
 481 GGCCCCGCGCACGGTGGGAAAGTTGGCCTGGAACCGGCCCGACCAGTTCCGCCCGGGCGCGAGGACCGGCCGCAGAAGTTGCTGCAAAACTTTTTGGGGGGTGCAGCCCGTGCCCCCCG

                                                                                                                    MetAlaMetValThrGly
 601 CGCGCCGGGGCCGAATGCGCGCCGCGTAGGGTCCCCCGGGCCGAGAGGGGTGCCCCGGAGGAAGAGCGCGGTGGGGGCGCCCCGGCCCCGCTGCCCTGGGGCTATGGCCATGGTGACCGGC

   GlyTrpGlyGlyProGlyGlyAspThrAsnGlyValAspLysAlaGlyGlyTyrProArgAlaAlaGluAspAspSerAlaSerProProGlyGlyAlaAlaSerAspAlaGluProGlyAsp
 721 GGCTGGGGCGGCCCCGGCGGCGACACGAACGGCGTGGACAAGGCGGGCGGCTACCCGCGCGCGGCCGAGGACGACTCGGCCTCGCCCCCCGGTGCCGCCAGCGACGCCGAGCCGGCGAC

   GluGluArgProGlyLeuGlnValAspCysValValCysGlyAspLysSerSerGlyLysHisTyrGlyValPheThrCysGluGlyCysLysSerPhePheLysArgThrIleArgArg
 841 GAGGAGCGGCCGGGGCTGCAGGTGGACTGCGTGGTGTGCGGGGACAAGTCGAGCGGCAAGCATTACGGTGTCTTCACCTGCGAGGGCTGCAAGAGCTTTTTCAAGCGAACGATCCGCCGC

   AsnLeuSerTyrThrCysArgSerAsnArgAspCysGlnIleAspGlnHisHisArgAsnGlnCysGlnTyrCysArgLeuLysLysCysPheArgValGlyMetArgLysGluAlaVal
 961 AACCTCAGCTACACCTGCCGGTCCAACCGTGACTGCCAGATCGACCAGCACCACCGGAACCAGTGCCAGTACTGCCGTCTCAAGAAGTGCTTCCGGGTGGGCATGAGGAAGGAGGCGGTG

   GlnArgGlyArgIleProHisSerLeuProGlyAlaValAlaAlaSerSerGlySerProProGlySerAlaLeuAlaAlaValAlaSerGlyGlyAspLeuPheProGlyGlnProVal
1081 CAGCGCGGCCGCATCCCGCACTCGCTGCCTGGTGCCGTGGCCGCCTCCTCGGGCAGCCCCCCGGGCTCGGCGCTGGCGGCAGTGGCGAGCGGCGGAGACCTCTTCCCGGGGCAGCCGGTG

   SerGluLeuIleAlaGlnLeuLeuArgAlaGluProTyrProAlaAlaAlaGlyArgPheGlyAlaGlyGlyGlyAlaAlaGlyAlaValLeuGlyIleAspAsnValCysGluLeuAla
1201 TCCGAACTGATCGCGCAGCTGCTGCGCGCTGAGCCCTACCCTGCGGCGGCCGGACGCTTCGGCGCAGGGGGCGGCGCGGCGGGCGCGGTGCTGGGCATCGACAACGTGTGCGAGCTGGCG

   AlaArgLeuLeuPheSerThrValGluTrpAlaArgHisGlyPhePheProGluLeuProValAlaAspGlnValAlaLeuLeuArgMetSerTrpSerGluLeuPheValLeuAsnAla
1321 GCGCGGCTGCTCTTCAGCACCGTGGAGTGGGCGCGCCACGGCTTCTTCCCCGAGCTGCCGGTGGCCGACCAGGTGGCGCTGCTGCGCATGAGCTGCGGAGCGAGCTCTTCGTGCTGAACGCG

   AlaGlnAlaAlaLeuProLeuHisThrAlaProLeuLeuAlaAlaAlaGlyLeuHisAlaAlaProMetAlaAlaGluArgAlaValAlaPheMetAspGlnValArgAlaPheGlnGlu
1441 GCGCAGGCGGCGCTGCCCCTGCACACGGCGCCGCTACTGGCCGCCGCCGGCCTCCACGCCGCGCCTATGGCCGCCGAGCGCGCCGTGGCTTTCATGGACCAGGTGCGCGCCTTCCAGGAG

   GlnValAspLysLeuGlyArgLeuGlnValAspSerAlaGluTyrGlyCysLeuAlaIleAlaLeuPheThrProAsnAlaCysGlyLeuSerAspProAlaHisValGluSerLeu
1561 CAGGTGGACAAGCTGGGCCGCCTGCAGGTCGACTCGGCCGAGTATGGCTGCCTCAAGGCCATCGCGCTCTTCACGCCCGACGCCTGTGGCCTCTCAGACCCGGCCCACGTTGAGAGCCTG

   GlnGluLysAlaGlnValAlaLeuThrGluTyrValArgAlaGlnTyrProSerGlnProGlnArgPheGlyArgLeuLeuArgLeuProAlaLeuArgAlaValProAlaSerLeu
1681 CAGGAGAAGGCGCAGGTGGCCCTCACCGAGTATGTGCGGGCGCAGTACCCGTCCCAGCCCCAGCGCTTCGGGCGCCTGCTGCTGCGGCTCCCCGCCCTGCGCGCGGTCCCTGCCTCCCTC

   IleSerGlnLeuPhePheMetArgLeuValGlyLysThrProIleGluThrLeuIleArgAspMetLeuLeuSerGlySerThrPheAsnTrpProTyrGlySerGlyGlnEnd
1801 ATCTCCCAGCTGTTCTTCATGCGCCTGGTGGGGAAGACGCCCATTGAGACACTGATCAGAGACATGCTGCTGTCGGGGAGTACCTTCAACTGGCCCTACGGCTCGGGCCAGTGACCATGA

1921 CGGGGCCACGTGTGCTGTGGCCAGGCCTGCAGACAGACCTCAAGGGACAGGGAATGCTGAGGCCTCGAGGGGCCTCCCGGGGCCCAGGACTCTGGCTTCTCTCCTCAGACTTCTATTTTT
2041 TAAAGACTGTGAAATGTTTGTCTTTTCTGTTTTTTAAATGATCATGAAACCAAAAAGAGACTGATCATCCAGGCCTCAGCCTCATCCTCCCCAGGACCCCTGTCCAGGATGGAGGGTCCA
2161 ATCCTAGGACAGCCTTGTTCCTCAGCACCCCTAGCATGAACTTGTGGGATGGTGGGGTTGGCTTCCCTGGCATGATGGACAAAGGCCTGGCGTCGGCCAGAGGGCTGCTCCAGTGGGCA
2281 GGGGTAGCTAGCGTGTGCCAGGCAGATCCTCTGGACACGTAACCTATGTCAGACACTACATGATGACTCAAGGCCAATAATAAAGACATTTCCTCACCTGC  2380
```

b

```
   1 CGCCGCCTGTGCCATTTCTGATTTTGCAACTTGGGGAAGAAGAAAAAAGCGAGAGAAGGGGAGCTTGCTCGCCGGGGGGTGGGGAGGGGGGAAGGAGACGCGGCCCCCCCAGGAACGGAG
 121 CGCGGGGGGAGCGGGCGAGGGGAGCAGGGGTGTTGGGGGGGGAGCCTGAGAGCCTGGGGGGGCTGCAAAAAGAGACAAAGAAAACAGCAGGAACCACAACAAAACGCCAGCAGGGCGGGC
 241 GGGCGCGCAGCAGCAGCGGGGCGGCCGAGGCAGTAGCGGCGGCAGCGGCGGCGGCGGAGGCAGCGGCCGGTGTCCGGCTCGGGCTCGGCTCCTGCGACCCCGGGGCGCCCGGCGGGC

                                                                                                                    Met
 361 CCCCCGCCCCCTCCCCCTCCCCCCTTCCCCTTCCCCTTCCCCTCCCAGCGCGCCCCGCGCGCCCCGCGGCCCTCGGCGAGCAGCTCGGCTCCCCCCAGCGCTCCCCGGGCCCAAAGATATG

   AlaMetValValSerSerTrpArgAspProGlnAspAspValAlaGlyGlyAsnProGlyGlyProAsnProAlaAlaGlnAlaAlaArgGlyGlyGlyGlyGlyAlaGlyGluGlnGln
 481 GCAATGGTAGTTAGCAGCTGGCGAGATCCGCAGGACGACGTGGCCGGGGGCAACCCCGGCGGCCCCAACCCCGCAGCGCAGGCGGCCCCGCGGCGGCGGCGCCGGCGAGCAGCAG

   GlnGlnAlaGlySerGlyAlaProHisThrProGlnThrProGlyGlnProGlyAlaProAlaThrProGlyThrAlaGlyAspLysGlyGlnGlyProProGlySerGlyGlnSerGln
 601 CAGCAGGCGGGCTCGGGCGCGCCGCACACGCCGCAGACCCCGGGCCAGCCCGGAGCGCCCGCCACCCCCGGCACGGCGGGGGACAAGGGCCAGGGCCCGCCCGGTTCGGGCCAGAGCCAG

   GlnHisIleGluCysValValCysGlyAspLysSerSerGlyLysHisTyrGlyGlnPheThrCysGluGlyCysLysSerPhePheLysArgSerValArgArgAsnLeuThrTyrThr
 721 CAGCACATCGAGTGCGTGGTGTGCGGGGACAAGTCGAGCGGCAAGCACTACGGCCAATTCACCTGCGAGGGCTGCAAAAGTTTCTTCAAGAGGAGCGTCCGCAGGAACTTAACTTACACA

   CysArgAlaAsnArgAsnCysProIleAspGlnHisHisArgAsnGlnCysGlnTyrCysArgLeuLysLysCysLeuLysValGlyMetArgArgGluAlaValGlnArgGlyArgMet
 841 TGCCGTGCCAACAGGAACTGTCCCATCGACCAGCACCACCGCAACCAGTGCCAATACTGCCGCCTCAAGAAGTGCCTCAAAGTGGGCATGAGGCGGGAAGCGGTTCAGCGAGGAAGAATG

   ProProThrGlnProAsnProGlyGlnTyrAlaLeuThrAsnGlyAspProLeuAsnGlyHisCysTyrLeuSerGlyTyrIleSerLeuLeuArgAlaGluProTyrProThrSer
 961 CCTCCAACCCAGCCCAATCCAGGCCAGTACGCACTCACCAACGGGGACCCCCTCAACGGCCACTGCTACCTGTCCGGCTACATCTCGCTGCTGCGCGCCGAGCCCTACCCACGTCG

   ArgTyrGlySerGlnCysMetGlnProAsnAsnIleMetGlyIleGluAsnIleCysGluLeuAlaAlaArgLeuLeuPheSerAlaValGluTrpAlaArgAsnIleProPhePhePro
1081 CGCTACGGCAGCCAGTGCATGCAGCCCAACAACATTATGGGCATCGAGAACATCTGCGAGCTGGCCGCGCGCCTGCTCTTCAGCGCCGTCGAGTGGGCCCGCAACATCCCCTTCTTCCCG

   AspLeuGlnIleThrAspGlnValSerLeuLeuArgLeuThrTrpSerGluLeuPheValLeuAsnAlaAlaGlnCysSerMetProLeuHisValAlaProLeuLeuAlaAlaAlaGly
1201 GATCTGCAGATCACCGACCAGGTGTCCCTGCTACGCCTCACCTGGAGCGAGCTGTTCGTGCTCAACGCGGCCCAGTGCTCTATGCCGCTGCACGTGGCGCCGTTGCTGGCCGCCGCCGGC

   LeuHisAlaSerProMetSerAlaAspArgValAlaPheMetAspHisIleArgIlePheGlnGluGlnValGluLysLeuLysAlaLeuHisValAlaSpSerAlaGluTyrSerCys
1321 CTGCATGCCTCGCCCATGTCTGCCGACCGCGTCGTGGCCTTCATGGACCACATCCGCATCTTCCAGGAGCAGGTGGAGAAGCTCAAGGCGCTACACGTCGACTCAGCCGAGTACAGCTGC

   LeuLysAlaIleValLeuPheThrSerAspAlaCysGlyLeuSerAspAlaAlaHisIleGluSerLeuGlnGluLysSerGlnCysAlaLeuGluGluTyrValArgSerGlnTyrPro
1441 CTCAAAGCCATCGTGCTGTTCACGTCAGACGCCTGTGGCCTGTCGGATGCGGCCCACATCGAGAGCCTGCAGGAGAAGTCGCAGTGCGCACTGGAGGAGTACGTGAGGAGCCAGTACCCC

   AsnGlnProSerArgPheGlyLysLeuLeuArgLeuProSerLeuArgThrValSerSerSerValIleGluGlnLeuPheHeValArgLeuValGlyLysThrProIleGlu
1561 AACCAGCCCAGCCGTTTTGGCAAACTGCTGCTGCGACTGCCCTCGCTGCGCACCGTGTCCTCCTCCGTCATCGAGCAGCTCTTCTTCGTCCGTTTGGTAGGTAAAACCCCCATCGAAACT

   LeuIleArgAspMetLeuLeuSerGlySerSerPheAsnTrpProTyrMetSerIleGlnCysSerEnd
1681 CTCATCCGCGATATGTTACTGTCTGGGAGCAGCTTCAACTGGCCTTACATGTCCATCCAGTGCTCCTAGACCTTGGGCGCTTCCCACCTGCCCCGTCCCCCTAGAGACTCAGAGGACCCA

1801 CCTGGGCCAAGGACTCCAAAGCCGCGGGGACACCGGGAAGTGCAGCGGGCCAGGCAGGCTGGGTGGGAGGGAGGAGGGCCGAGACAGGAGCAGCCCACCCAGCAGAAATACAATCCGAGC
1921 TACAAAGCATGGGAAAAAGAGACTCTTTAGGATCAGATCTGTGAGCACGTTGGCGAGGAAAAACAAACAAAAAAAGAACCTTGTGTCTGTCTGGTGAAAAAAGAAAAAGCAAA
2041 TTGGAAGAGAGGACCATGAGAATTTTAATAAAACAGAAGGAAACTAATGGACCTTCCAGGATTTATTGTGGACGGATGTGGATATATTCTGTACAGGAACAACACATATGGAAGTGGACT
2161 GAAGCCTATGTAGAAACACACACACACTGAACATTGTTATTCATTTTGTAAAATACTAGTCTTTATTTTCATTTTTTGTAAAATTTAAACATCGTATGCGCATAAAGAAAAAGGAAACAA
2281 GAATTAGGGGAAAATAACATTTTCCAAATAATTATAAAAAATTGTCCTGTGTCTATGTATCTATATCTGTTTTGTATTTTTTTCTGGTTCCAAACCAGATTTCCTGTGATTCTATACTAA
2401 TAATTTTTGATATAACCCTTTGCTTCTTATAATGAGTGCGATATATGTTGTCGAGGCTGTTCTTCAAGAATTAAAATTGAAGTGAAAATTTAAACAAAAATAAAAGAATTG  2511
```

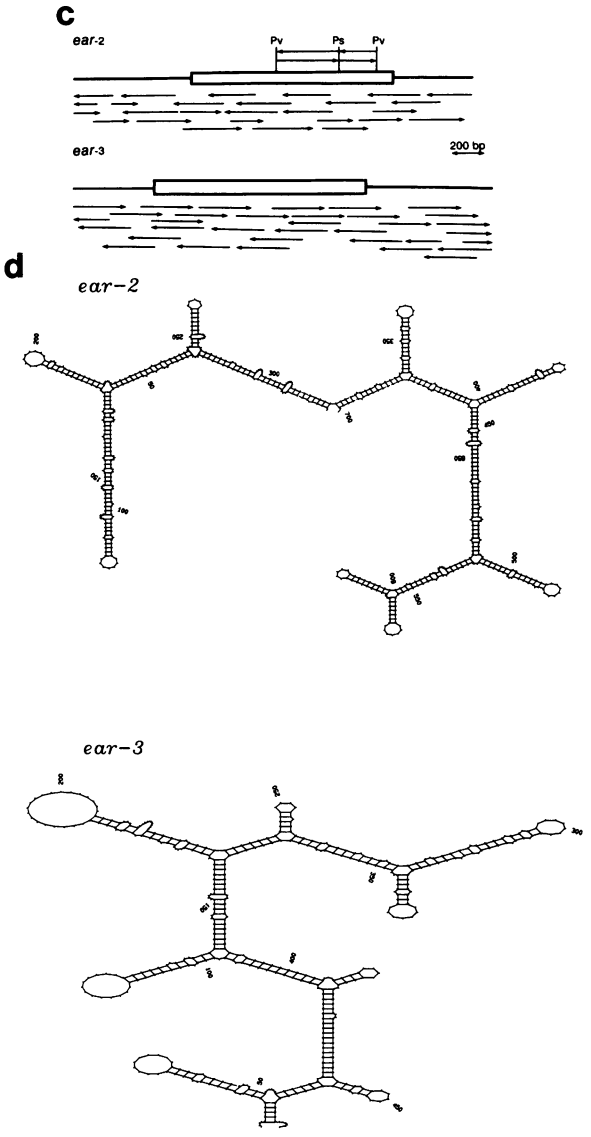

Figure 4. Nucleotide sequences of the ear-2 and ear-3 cDNAs. The nucleotide sequences of the cDNAs for the ear-2 gene (a) and ear-3 gene (b) were determined by the dideoxy chain termination method. The deduced amino acid sequences of the ear proteins are shown above the corresponding nucleotide sequences. The sequencing strategy is shown in (c). The direction and extent of sequence determination are shown by horizontal arrows. The inserts of λA14 (for ear-2) and λA3

(for ear-3) were successively deleted from either end by the method of Henikoff (39) and the resulting DNAs were subjected to sequencing (see MATERIALS and METHODS). From the sequence of λA14 insert, the unexpectedly small-sized protein product of ear-2 was predicted, which suggested some deletion in the insert. Therefore we recloned a 600-bp PvuII fragment of λA15 into pUC18 and pUC19 plasmids and sequenced the 600-bp fragment as shown above the schematic illustration of λA14 insert. The open boxes illustrated in the DNA structures of λA14 and λA3 inserts represent protein-coding regions.(d) Secondary structures of the 5' untranslated region of the ear2 and ear3 mRNAs were derived by the program FOLD and drawn by SQUIGGLES (41,42). G-C,A-T and G-T pairs are indicated by solid bars.

derivative chromosome 20, which was prepared from cell lines carrying reciprocal translocation t(1;20) (data not shown).

Next we examined the chromosomal location of ear-3. The ^{32}P-labeled ear-3 probe, a 2.3-kbp EcoRI insert of λA11 (Fig. 2), hybridized with DNA in fraction "B", which consisted of chromosomes 3, 4 and 5 of GM0131 (Fig. 6A-b). Then we carried out hybridization experiments with DNAs prepared from the sorted

Figure 5. Comparison of the predicted amino acid sequences of the ear-2 and ear-3 gene products. Amino acids of the two proteins are numbered along the sequence. Identities in the sequences are marked by two dots between the two lines; predicted DNA binding domains are boxed.

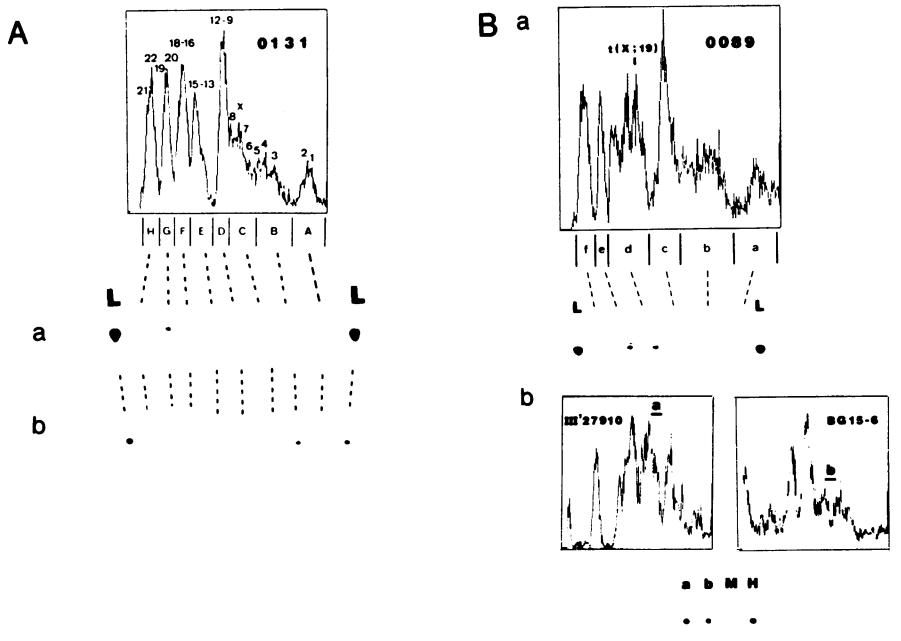

Figure 6. Southern hybridization of DNAs from sorted chromosomes.
(A) Chromosomes of lymphoblast GM0131 cells were sorted into 8 fractions (A to H). DNA from each fraction was cleaved with EcoRI and subjected to Southern blot analysis with the ^{32}P-labeled ear-2 (a) and ear-3 (b) probes. (B) Chromosomes of fibroblast GM0089 cells (a) and human-mouse hybrid cells III' 27910 and BG15-6 (b) (kindly donated by Y. Kaneda, Osaka University, Osaka) which contain human chromosome 5 alone, were sorted into fractions (a , fractions a to f; b , a fraction containing the human chromosome). DNAs from the fractions were analyzed with ear-2 (a) and ear-3 (b) probes, after their cleavage with EcoRI (a) or PvuII (b). Total human DNA (Lanes L and H) or mouse DNA (Lane M) cleaved with EcoRI (a) or PvuII (b) was used as a standard.

chromosomes of human-mouse hybrid-cells that contained chromosome 3, chromosome 4 or chromosome 5 of human origin besides the mouse chromosomes. The ear-3 probe did not hybridize with DNAs prepared from the hybrids that contained human chromosome 3 or 4 (data not shown). However, the ear-3 probe did hybridize with DNAs from cell lines III'27910 and BG15-6 that contained human chromosome 5 (Fig. 6B-b), indicating that the ear-3 gene was located on chromosome 5.

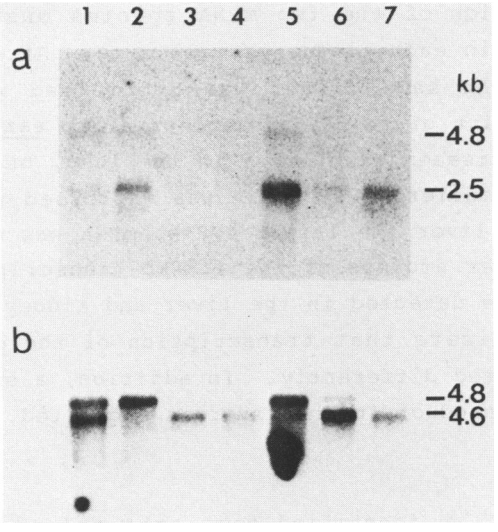

Figure 7. Blot hybridization of RNAs from various cell types with ear probes.
Samples (2µg) of poly(A)$^+$ RNA prepared from human embryo fibroblasts (lane 1), from A431 squamous carcinoma cells (lane 2), from the lung (lane 3), brain (lane 4), liver (lane 5) and kidney (lane 6) of a human fetus and from placenta at term (lane 7) were subjected to blot hybridization with the ^{32}P-labeled ear-2 probe (a) and ear-3 probe (b) under stringent conditions. Hybridizations with the second probe were carried out after decay of the radioactivity of the previously hybridizing probes. The length of each transcript was determined by comparison with those of transcripts of the human EGF receptor gene (10.5-, 5.6-, and 2.9-kb), which were seen in lane 2 of the same filter hybridized with the EGF receptor probe, pE7 insert DNA (40).

Expression of ear-2 and ear-3 in normal tissues.

Northern hybridization of RNAs prepared from human fetal tissues (lung, brain, liver and kidney), placenta, cultured embryo fibroblasts and A431 squamous carcinoma cells was carried out with each ear-specific probe under stringent conditions (Fig. 7). The 2.5-kb ear-2 mRNA was expressed at a higher level in the liver than in other tissues. The 2.5-kb ear-2 transcripts were hard to detect in the lung and brain. Minor amounts of the 4.8-kb transcripts were observed in the liver, embryo fibroblasts and A431 cells but not in the other tissues. Since the amount of the 2.5-kb mRNA relative to the 4.8-kb mRNA

varies, expression of the two mRNA species must be regulated differentially in each tissue or cell type. After decay of the first probe, the same filter was hybridized with the ear-3 probe. Expression of the 4.6-kb and/or 4.8kb ear-3 mRNA was observed in all tissues examined. In the lung, brain, kidney and placenta, the smaller 4.6-kb mRNA was expressed predominantly, whereas in the liver the larger 4.8-kb mRNA was expressed most. Relatively higher amounts of the 4.8-kb transcripts and 4.6-kb transcripts were detected in the liver and kidney, respectively. These data indicate that transcription of the ear-2 and ear-3 genes is regulated differently. In addition, a specific role of the ear-2 gene product in the liver is suggested.

DISCUSSION

In this study, we identified four erbA-related genes, named ear-1, ear-2, ear-3 and ear-7, and characterized the ear-2 and ear-3 cDNA clones. The identification of several erbA homologs in the human genome has been reported by various groups. These homologs were named erbA1 and erbA2 (30), and erbA (17). The erbA1 gene, which was mapped on chromosome 17q21.3, showed the highest homology with the v-erbA gene (14). Since the ear-7 gene product also showed very high homology, not only in the DNA binding domain (Fig. 3) but also in the putative hormone binding domain (to be published), with the v-erbA gene product, we assume that the ear-7 gene is the same as the erbA1 gene. The existence of other erbA-related genes was also demonstrated by low-stringency hybridization experiments using an erbA specific probe or erbA2 specific probe (16,17). In addition, a human erbA-related gene was shown to be located close to the integration site of a hepatitis virus B genome (33). This gene was recently shown to encode a receptor protein for retinoic acid (18,19). Since none of these genes were mapped on chromosome 5 or chromosome 19, they are apparently different from the ear-2 and ear-3 genes. We have recently mapped both the ear-7 and ear-1 genes on human chromosome 17 and found that the two genes were closely linked (to be published). Thus, cellular erbA homologs form a superfamily consisting of at least 9 genes: erbA1/ear-7, erbA-2 (probably the same gene as the ear-1 gene),

Table 1. Homologies of amino acid sequences of the ear-2 protein and other products of the erbA-related genes.

% Homology with ear-2	ear-1	ear-3	ear-7	v-erbA	$T_3R\alpha$	$T_3R\beta$
DNA binding region	57	84	60	62	60	62
Ligand binding region	23	70	<15	21	21	20
	GR	MR	PR	ER	RAR	HAP
DNA binding region	51	53	49	53	60	60
Ligand binding region	<15	<15	<15	<15	22	22

Values are percentage of identical amino acids in the specific regions of the protein products of the erbA-related genes. Homologies were examined in the DNA binding region, from amino acids 56 to 121 of ear-2 and corresponding regions of other proteins, and in the ligand binding region, from amino acids 184 to 403 of ear-2 and corresponding regions of other proteins. The two regions of v-erbA, T_3R(T_3-receptor)α, $T_3R\beta$, GR(glucocorticoid receptor), MR(mineralcorticoid receptor), RAR(retinoic acid receptor), and HAP(a protein product encoded by a cellular gene at a site of hepatitis B virus integration in a hepatocellular carcinoma) were referred from a review by Evans (34). The DNA binding regions and ligand binding regions of ear-1 and ear-7 were defined from our unpublished data.

erbA2-homolog at chromosome 17q25, erbA , two erbA-homologs at chromosome 3 (one of which may correspond to the erbA-homolog at the hepatitis B integration site), retinoic acid receptor gene, ear-2, and ear-3. In addition, genes encoding receptors for adrenal- and sex-steroids were also shown to be members of this family (34). Southern blot hybridization of chicken and mouse genomic DNAs with ear specific probes showed that the ear-1, ear-2, ear-3 and ear-7 genes were independently conserved in these animals (data not shown).

Steroids and the thyroid hormone exert their effects via essentially similar mechanisms. Namely they enter cells, possibly by passive diffusion, and bind to high affinity nuclear receptors that in turn bind to nuclear DNA sequences resulting in selective activation of gene expression (35,36). Accordingly, the receptors ought to carry two functional domains, the DNA

binding domain and the ligand binding domain. Thus, it is reasonable to predict that the ear-2 and ear-3 gene products, whose structures are characteristic in that they consist of the DNA binding domain and the possible ligand binding domain, are members of the receptor for ligands that are biologically and structurally similar to steroid hormones. However, comparison of the known sequences of the steroid hormone receptors and sequences of the ear-2 and ear-3 proteins revealed that ear-2 and ear-3 are distinct from the previously published genes encoding steroid hormone receptors (Table 1). Polycyclic aromatic hydro-carbons, such as 3-methylcholanthrene and dioxin derivatives, which have been postulated to regulate expression of the gene encoding cytochrome P-450 by binding to the high affinity receptors (37), could also be ligands of the erbA-homologs.

Recently, the avian c-erbA gene has been shown to encode the receptor for thyroid hormone T_3 (7), indicating that the human counterpart of this gene, erbA1, encodes a T_3-receptor. Interestingly, however, the erbAβ gene was also shown to encode a T_3-receptor (17) with a high affinity similar to that of avian c-erbA protein. Similarly, two apparently distinct but highly related receptor molecules, the ear-2 and ear-3 proteins (Fig. 5 and Table 1), would have been formed by a recent split within the erbA gene superfamily and could react with a same ligand. In addition, the predicted DNA binding domains of the ear proteins are rich in cysteines and have sequence homology with that proposed for Xenopus 5S gene transcription factor IIIA. The high homology between the DNA binding domains of the ear-2 and ear-3 proteins (57/66 match the amino acid sequences) suggest that the two proteins recognize similar regulatory sequences, as in the case of the two related proteins, glucocorticoid receptor and mineralcorticoid receptor, when they activate MMTV(mouse mammary tumor virus) LTR(long terminal repeat) transcription in response to dexamethasone (38).

Since the ear-2 gene is expressed at relatively high levels in fetal liver, the ear-2 protein may control the expression of genes relevant to hematopoietic cells. More detailed northern blot analyses of RNA from various types of cells are needed to

understand the functions of ear-2 and ear-3 proteins. Expression of a full-length cDNA under LTR transcriptional control in cultured cells would be helpful in examining the functions of these erbA-related gene products.

ACKNOWLEDGEMENTS

We thank Y. Yamamuro for her help in preparation of the manuscript. This work was supported in part by a Grant-in-Aid for Special Project Research Cancer-Bioscience from the Ministry of Education, Science, and Culture of Japan.

REFERENCES

1. Russel M, Saule, S., Lagrou, C., Beug, H., Graf, T. and Stehelin, D. (1979) Nature 281, 452-455
2. Vennstrom, B., and Bishop, J. M. (1982) Cell 28, 135-143
3. Yamamoto, T., Hihara, H., Nishida, T., Kawai, S., and Toyoshima, K. (1983a) Cell 34, 225-232
4. Frykberg, L., Palmieri, S., Berg, H., Graf, T., Hayman, M. J., and Vennstrom, B. (1983) Cell 32, 227-238
5. Yamamoto, T., Nishida, T., Miyajima, N., Kawai, S., Ooi, T., and Toyoshima, K. (1983b) Cell 35, 71-78
6. Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. and Waterfield, M.D. (1984) Nature 307, 521-527
7. Sap, J., Munoz, A., Damm, K., Goldberg, Y., Ghysdael, J., Leutz, A., Berg, H., and Vennstrom, B. (1986) Nature 324, 635-640
8. Martial, J.A., Baxter, J.D., Goodman, H.M. and Seeburg, P.H. (1977) Proc. Natl. Acad. Sci. USA 74, 1816-1820
9. Semba, K., Kamata, N., Toyoshima, K., and Yamamoto, T. (1985) Proc. Natl. Acad. Sci. USA 82, 6497-6501
10. Yamamoto, T., Ikawa, S., Akiyama, T., Semba, K., Nomura, N., Miyajima, N., Saito, T., Toyoshima, K. (1986) Nature 319, 230-234
11. Bargmann, C., Hung, M.-C. and Weinberg, R.A. (1986) Nature 319, 226-230
12. Coussens, L., Yang-Feng, T.L., Liao, Y.-C., Chen, E., Gray, A., McGrath, J., Seeburg, P.H., Libermann, T. A., Schlessinger, J., Francke, U., Levinson, A., and Ullrich, A. (1985) Science 230, 1132-1139
13. Sheer, D., Sheppard, D.M., Lebear, M., Rowley J.D., San Roman, C., and Salomon, E. (1985) Ann. Hum. Genet. 49, 167-172
14. Dayton, A.I., Selden, J.R., Laws, G., Dorney, D.J., Finan, J., Tripputi, P., Emanuel, B.S., Rovera, G., Nowell, P.C. and Croce, C.M. (1984) Proc. Natl. Acad. Sci. USA 81, 4495-4499
15. Spurr, N.K., Solomon, E., Jansson, M., Sheer, D., Goodfellow, P.N., Bodmer, W.F., Vennstrom, B. (1984) EMBO, J. 3, 159-163
16. Gosden, J.R., Middleton, P.G., Rout, D., and DeAngelis, C. (1986) Cell Genet. 43, 150-153

17. Weinberger, C., Thompson, C.C., Ong, E.S., Lebo, R., Gruol, D.J., and Evans, R.M. (1986) Nature 324, 641-646
18. Benbrook, D., Lernhardt, E., and Pfahl, M. (1988) Nature 333, 669-672
19. Petkovich, M., Brand, N. J., Krust, A. and Chambon, P. (1987) Nature 330, 444-450
20. Ohashi, M., Aizawa, S., Ooka, H., Ohsawa, T., Kaji, K., Kondo, H., Kobayashi, T., Nomura, T., Matsuo, M., Mitsui, Y., Murota, S., Yamamoto, K., Itoh, H., Shimada, H., and Utakoji, T. (1980) Exp. Gerontol. 15, 121-133
21. Huynh, T.V., Young, R.A. and Davis, R.W. (1985) In DNA Cloning: A Practical Approach (ed. Glover, D.) I, 4978 (IRL Press, Oxford)
22. Messing, J. (1983) Methods Enzymol 101, 20-78
23. Mizusawa, S., Nishimura, S. and Seela, F. (1986) Nucleic Acids Res. 14, 1319-1324
24. Hattori, M. and Sasaki, Y. (1986) Anal. Biochem. 152, 232-238
25. Chirgwin, J.M., Pyzybyla, A.E., McDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294-5299
26. Lehrach, H., Diamond, K., Wozney, J.M. and Boedtken, H. (1977) Biochemistry 16, 4743-4751
27. Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205
28. Fukushige, S., Murotsu, T. and Matsubara, K. (1986) Biochem. Biophys. Res. Commun. 134, 477-483
29. Nishida, T., Sakamoto, S., Yamamoto, T., Hayman, M., Kawai, S. and Toyoshima, K. (1984) Gann 75, 325-333
30. Jansson, M., Philipson, L. and Vennstrom, B. (1983) EMBO, J. 2, 561-565
31. Kozak, M. (1984) Nature 308, 241-246
32. Frier, S.M., Kierzek, R., Jager, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T and Turner, D.H. (1986) Proc. Natl. Acad. Sci. USA 83, 9373-9377
33. Dejean, A., Bougueteret, L., Grzeshik, K.-H. and P. Tiollais (1986) Nature 322, 70-72
34. Evans, R.M. (1988) Science 240, 889-895
35. Macleod. K.M. and Baxter, J.D. (1976) J. Biol. Chem. 251, 7380-7387
36. Yamamoto, K.R., (1985) Ann. Rev. Genet. 19, 209-252
37. Poland, A. and Glover, E. (1976) J. Biol. Chem. 251, 4936-4946
38. Arriza, J.L., Weinberger, C., Cerlli, G., Glaser, T.M., Handelin, B.L., Housman, D.E. and Evans, R.M. (1987) Science 237, 268-275
39. Henikoff, S. (1984) Gene 28, 351-359
40. Merlino, G.T., Xu, Y.H., Ishii, S., Clark, A.J.L., Semba, K., Toyoshima, K., Yamamoto, T., and Pastan, I. (1984) Science 224, 417-419
41. Zuker, M. (1981) Nucleic Acids Res. 9, 133-148
42. Devereux, J., Haeberli, P and Smithies, O. (1984) Nucleic Acids Res. 12 387-395