SHORT COMMUNICATION

The Epimorphin Gene Is Highly Conserved among Humans, Mice, and Rats and Maps to Human Chromosome 7, Mouse Chromosome 5, and Rat Chromosome 12

HONGBIN ZHA, ELAINE F. REMMERS, CLAUDE SZPIRER,* JOSIANE SZPIRER,* HEYING ZHANG, CHRISTINE A. KOZAK,† AND RONALD L. WILDER[1]

*Arthritis and Rheumatism Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland 20892; *Universite Libre de Bruxelles, Departement de Biologie Moleculaire, Rue des Chevaux, 67, B-1640 Rhode-St-Genese, Belgium; and †Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892*

Received September 29, 1995; accepted August 30, 1996

A genomic DNA fragment containing the rat epimorphin gene sequence was cloned from a rat DNA cosmid library using a mouse epimorphin cDNA probe. Within the cosmid insert, nine epimorphin exons were identified and sequenced. The predicted amino acid sequence of the rat epimorphin protein exhibited 96 and 86% identity with the mouse and human epimorphin proteins, respectively. Consistent with the developmentally related expression pattern of the mouse epimorphin gene, transcripts of the rat epimorphin gene were detected in 17-day postfertilization rat embryos. The gene, designated *Epim*, was assigned to rat chromosome 12 by somatic cell hybrid analysis and localized to 12q16 by fluorescence *in situ* hybridization. The mouse and human homologs of this gene were localized on mouse chromosome 5 and human chromosome 7 by linkage analysis and chromosomal *in situ* hybridization, respectively.

Epimorphin is a protein expressed on the surface of mouse embryonic mesenchymal cells (6). The human homolog of this gene has also been described (7). This protein is involved in epithelial morphogenesis, including hair follicle growth and lung epithelial tubule formation. Here we report the genetic cloning of rat epimorphin and the chromosomal mapping of this gene in rats, mice, and humans.

A mouse epimorphin cDNA probe was prepared from total RNA extracted from C3H10T1/2 cells (American Type Culture Collection, Rockville, MD). These cells were previously shown to express the epimorphin gene when grown to confluence (6). Total RNA (5 μg) was reverse transcribed in a 50-μl reaction containing 50 m*M* Tris–HCl (pH 8.3), 40 m*M* KCl, 6 m*M* $MgCl_2$, 20 m*M* dithiothreitol, 0.4 m*M* each dATP, dCTP, dGTP, and dTTP, 2.4 ng/μl random hexamers (Perkin–Elmer/Cetus Corp., Foster City, CA), 0.5 U/μl RNase inhibitor (Perkin–Elmer Corp.) and 400 U MMLV reverse transcriptase (Life Technologies, Gaithersburg, MD). The reaction mix was incubated at 22°C for 10 min, and 42°C for 30 min and heated to 95°C for 10 min. The mouse epimorphin cDNA fragment was then amplified in a 50-μl reaction containing 6 μl of the cDNA mixture, 50 m*M* KCl, 10 m*M* Tris–HCl (pH 8.3), 2 m*M* $MgCl_2$, 200 μM each dATP, dCTP, dGTP, and dTTP, 4 μM each sense and antisense primers, and 2.5 U of *Taq* polymerase (Perkin–Elmer Corp.). The primers selected from the mouse epimorphin sequence reported by Hirai *et al.* (6) were sense, 5′-GTTTCTTCCATC-AGGTAGAGGAG-3′ and antisense, 5′-GCAATCATT-TGCCAACCGACAAG-3′. The 780-bp cDNA product was gel-purified using a QIAEX kit (Qiagen, Chatsworth, CA) and used as a hybridization probe.

A rat genomic cosmid library with inserts of 30–48 kb (Clontech Laboratories, Inc., Palo Alto, CA) was screened with the mouse cDNA probe labeled by random priming (Boehringer Mannheim, Indianapolis, IN). A single cosmid clone containing the rat epimorphin gene was detected and the cosmid DNA prepared. The purified cosmid DNA was sequenced using a fmol DNA Sequencing system (Promega, Madison, WI). Using primers designed from the mouse cDNA coding region or the newly obtained rat sequence, rat sequences were determined and compared to the mouse cDNA sequence. Intron/exon junctions were identified where the rat genomic sequence diverged from the mouse cDNA. In this way, nine contiguous exons were identified and sequenced. The rat genomic DNA sequences obtained were submitted to GenBank (Accession Nos. U35039 to U35047). Figure 1 shows the deduced amino acid sequence of the rat epimorphin protein as well as

Sequence data from this article have been deposited with the GenBank/EMBL Data Libraries under Accession Nos. U35039–U35047.

[1] To whom correspondence should be addressed at Building 10, Room 9N228, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892. Telephone: (301) 496-3373. Fax: (301) 402-0012.

GENOMICS **37,** 386–389 (1996)
ARTICLE NO. 0574
0888-7543/96 $18.00

```
Mouse     MRDRLPDLTA CRTNDDGDTA VVI-VEKDHF MDGFFHQVEE IRSSIARIAQ
                     **..****.* *.* ****** **.******* **********
Rat                  CRKSDDGDNA VIITVEKDHF MDAFFHQVEE IRSSIARIAQ
                     ***.****   ....****** **.******* **.**..*.*
Human     MRDRLPDLTA CRKNDDGD-- TVVVVEKDHF MDDFFHQVEE IRNSIDKITQ

          HVEDVKKNHS IILSAPNPEG KIKEELEDLD KEIKKTANRI RGKLKSIEQS
          ********** ********** *********. ********** *****.****
          HVEDVKKNHS IILSAPNPEG KIKEELEDLN KEIKKTANRI RGKLKAIEQS
          .**.****** ********** ********** ********.*  .********
          YVEEVKKNHS IILSAPNPEG KIKEELEDLN KEIKKTANKI AAKLKAIEQS

          CDQDENGNRT SVDLRIRRTQ HSVLSRKFVD VMTEYNEAQI LFRERSKGRI
          ********** ********** ********** ********** *****.****
          CDQDENGNRT SVDLRIRRTQ HSVLSRKFVD VMTEYNEAQI LFRERRKGRI
           ****.**** ********** *********. .*.******. *****.****
          FDQDESGNRT SVDLRIRRTQ HSVLSRKFVE AMAEYNEAQT LFRERSKGRI

          QRQLEITGRT TTDDELEEML ESGKPSIFIS DIISDSQITR QALNEIESRH
          ********** ***.****** ********** ********** **********
          QRQLEITGRT TTDEELEEML ESGKPSIFIS DIISDSQITR QALNEIESRH
          ********** ***.****** ********.* ********** **********
          QRQLEITGRT TTDDELEEML ESGKPSIFTS DIISDSQITR QALNEIESRH

          KDIMKLETSI RELHEMFMDM AMFVETQGEM VNNIERNVVN SVDYVEHAKE
          ********** ********** ********** ********** **********
          KDIMKLETSI RELHEMFMDM AMFVETQGEM VNNIERNVVN SVDYVEHAKE
          ********** ********** ********** .*******.* ..********
          KDIMKLETSI RELHEMFMDM AMFVETQGEM INNIERNVMN ATDYVEHAKE

          ETKKAIKYQS KARRKKWIIA AVAVAVIAVL ALIIGLSVGK
          ********** ********** **.******* **********
          ETKKAIKYQS KARRKKWIIA AVVVAVIAVL ALIIGLSVGK
          ********** *********  ** *..... .*********
          ETKKAIKYQS KARRKKWIII AVSVVLVVII VLIIGLSVGK
```

FIG. 1. Comparison of the predicted amino acid sequence of rat epimorphin with those of mouse and human epimorphin. Identical and conservative residues are marked with asterisks and dots, respectively. We were unable to identify the first 10 amino acids of the rat epimorphin.

alignments with the amino acid sequences of human (7) and mouse (6) epimorphin proteins. The primary structure of the rat epimorphin protein is well conserved among these species, with the rat sharing 96 and 86% identity with the mouse and human, respectively. Using mouse primers, we failed to identify the rat DNA sequence homologous to that encoding the first 10 amino acids of the mouse protein. It is possible that the nucleotide sequence in this region is not well conserved between the two species or is not included in the cloned genomic DNA fragment.

Mouse epimorphin is expressed in a developmentally related fashion. It can be detected in cultured 13-day mouse embryonic skin rudiments with vigorous outgrowth of whisker hair follicles (6). Similarly, we detected rat epimorphin mRNA in 17-day postfertilization Sprague–Dawley (SD/Hsd) rat embryos by nonspecific reverse transcription and specific PCR amplification as described above but using rat exon sequence primers: sense, 5′-GTGTAGGAAAAGCGACGATGGAGAC-3′; antisense, 5′-TCATTTGCCAACCGACAAGCCAATG-3′. The amplified rat cDNA fragment was gel purified and used as a hybridization probe for mapping the rat and human epimorphin genes.

The rat epimorphin gene was assigned to a rat chromosome by somatic cell hybrid analysis. The hybrids used in this study were derived from the fusion of mouse hepatoma cells (BWTG3) with adult rat hepatocytes (HSRD) (13). *Hin*dIII-digested DNAs from hybrid cell lines and nonhybrid rat and mouse cells were blotted onto nylon. The rat and mouse epimorphin DNA fragments were detected by hybridization with the rat cDNA probe (840 bp) labeled with [α-32P]dCTP by random priming (Boehringer Mannheim). A single hybridization band was detected from either rat or mouse DNA. The rat and mouse hybridization bands were distinct. Of 17 hybrid cell lines, 12 exhibited the specific rat DNA fragment. Comparison of these results with the rat chromosome content for each hybrid cell line indicated that the rat-specific fragment segregated

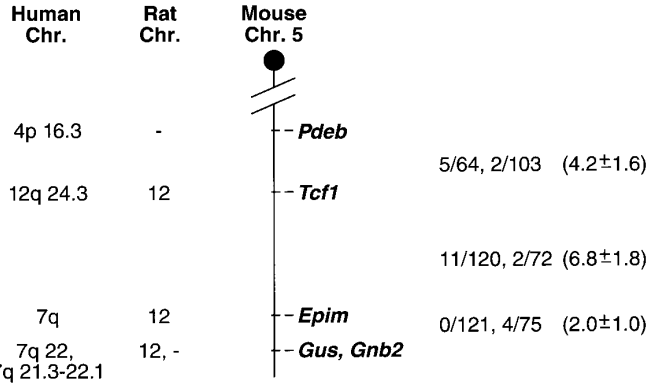

FIG. 2. An abbreviated map of mouse chromosome 5 showing the location of *Epim*. Recombination fractions are given to the right of the map for each adjacent locus pair, with the first fraction representing data from the *M. m. musculus* cross and the second fraction representing data from the *M. spretus* cross. Numbers in parentheses represent the percentage recombination and standard error calculated according to Green (5). The chromosomal locations of the human and rat homologous loci are indicated to the left of the mouse map.

with chromosome 12 with complete concordance. At least two discordant clones were found for each of the other chromosomes.

Regional chromosome localization of the rat epimorphin gene was obtained by fluorescence *in situ* hybridization (FISH) using as a probe the rat cosmid DNA described above labeled with biotin (11, 12). The probe generated bright double chromatid signals on chromosome 12 only. The position of double chromatid signals along the chromosome was measured on 15 chromosomes and found to be between 73 and 87% of the chromosome length (starting from the p arm telomere), allowing the assignment of the rat epimorphin gene to 12q16 (Fig. 3).

The mouse chromosomal location of the epimorphin gene was determined by linkage analysis using two genetic crosses: (NFS/N or C58/J × *Mus spretus*) × *Mus musculus musculus* (9) and (NFS/N × *M. spretus*) × *M. spretus* or C58/J (1). Southern blot hybridization using the mouse cDNA probe identified *Eco*RI fragments of 11.7 kb in *M. m. musculus* and 6.7 kb in NFS/N and C58/J. Digestion with *Bam*HI identified fragments of 17.0 and 9.9 kb in *M. spretus* and NFS/N, respectively. Inheritance of the polymorphic fragments in the progeny of the two crosses was compared with inheritance of over 750 markers previously mapped to all 19 autosomes and the X chromosome. As shown in Fig. 2, the gene encoding epimorphin, designated *Epim*, was mapped to chromosome 5 and showed closest linkage to the markers *Pdeb* (phosphodiesterase B), *Tcf1* (transcription factor 1 or *Hnf1*), *Gus* (J glucuronidase), and *Gnb2* (guanine nucleotide binding protein B-2). Probes for these markers and RFLPs used to type these crosses have been described

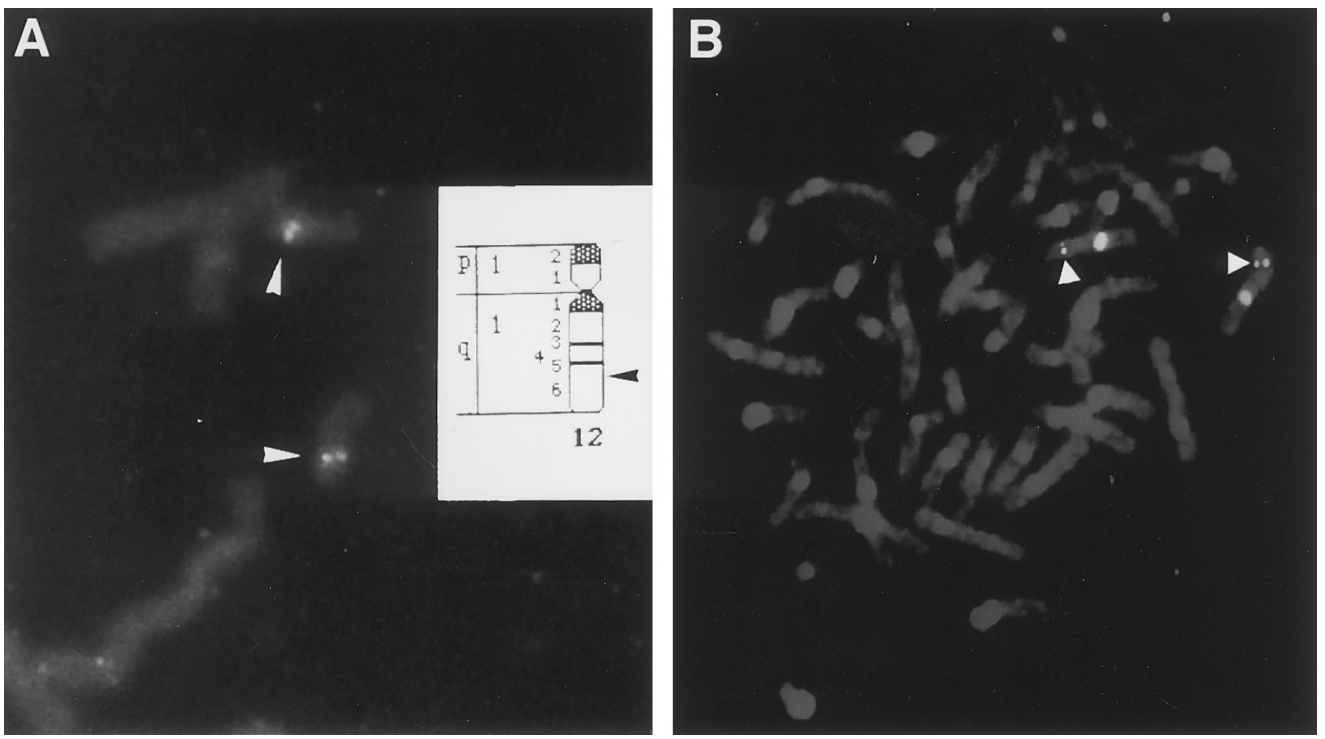

FIG. 3. The location of the epimorphin gene on rat and human chromosomes. (**A**) Regional localization of the rat epimorphin gene by FISH. A portion of a metaphase showing the signals (arrowheads) generated by the cosmid probe on the four chromatids of rat chromosome 12. The diagram (inset) illustrates the average position of the signals at 78% of the chromosome length, corresponding to 12q16. (**B**) Localization of the human epimorphin gene on human chromosome 7. The localization was performed by FISH using a biotin-labeled rat cDNA probe hybridized to human metaphase chromosomes. The strong signals are from a human chromosome 7 I satellite centromere-specific probe cohybridized with the rat cDNA *Epim* probe (weaker signals are indicated with arrowheads).

previously (2, 4). The epimorphin gene is thus a new member of a synteny group conserved between mouse chromosome 5 and rat chromosome 12, which also includes *Tcf1*, *Mdh2*, and *Gus* (8).

The human homolog of this gene was localized by FISH using a biotin-labeled rat cDNA probe hybridized to human metaphase chromosomes with subsequent signal detection by FITC (Paragon Biotech Inc., Baltimore, MD). Of 25 metaphase spreads, three showed a positive signal on the long arm of chromosome 7 (Fig. 3). Based on our FISH results and linkage homology data published previously between mouse and human (3), we provisionally assign the epimorphin gene (human locus name EPIM) to the long arm of human chromosome 7.

Epimorphin, originally identified in mouse embryonic tissues by Hirai and co-workers, was reported to be essential for epithelial cell organization. A monoclonal antibody against this protein perturbed the growth and reconstruction of hair follicles as well as the tubular formation of lung epithelial cells in organ culture (6). There is, however, a large discrepancy between the molecular weights of the epimorphin protein recognized by its monoclonal antibody (150 kDa) and the protein predicted by its cloned cDNA (34 kDa). Although a 150-kDa immunoreactive protein is detected in cells transfected with the cDNA, it is still possible that the cloned cDNA and the genomic DNA fragment do not represent the epimorphin protein detected by the monoclonal antibody. Recently, Pelham noted that the protein encoded by the mouse epimorphin cDNA cloned by Hirai and co-workers appears to be a member of a family of proteins involved in the targeting and/or fusion of intracellular vesicles to specific membranes (10). Determination of peptide sequences of gel-purified epimorphin protein should help resolve this controversy.

The map location of the mouse *Epim* gene places it in a region of chromosome 5 between regions of linkage homology to human chromosomes 7 and 12 (3). The location of this gene on human chromosome 7 thus extends proximally the region of homology on mouse chromosome 5 to human chromosome 7. The position of this gene also places it in close proximity to the developmental mutation *bf* (buff), characterized by alteration of the coat color in nonagouti mice. The involvement of epimorphin in hair follicle growth suggests that it may be a potential candidate for this developmental disorder.

ACKNOWLEDGMENTS

We thank Michele Riviere and Pascale Vanvooren for excellent technical assistance. Work in Brussels was supported by the CGER-Assurances and the Fund for Scientific Medical Research (FRSM). C.S. is a Research Director of the FNRS (Belgium).

REFERENCES

1. Adamson, M. C., Silver, J., and Kozak, C. A. (1991). The mouse homolog of the Gibbon ape leukemia virus receptor: Genetic mapping and a posssible receptor function in rodents. *Virology* **183:** 778–781.
2. Danciger, M., Farber, D. B., Peyser, M., and Kozak, C. A. (1990). The gene for the β-subunit of retinal transducin (*Gnb-1*) maps to distal chromosome 4, and related sequences map to mouse chromosomes 5 and 8. *Genomics* **6:** 428–435.
3. Debrey, R. W., and Seldin, M. F. (1996). Human/mouse homology relationships. National Center for Biotechnology Information. http://www3.ncbi.nlm.nih.gov/Homology/.
4. Gonzales F. J., Liu, S.-Y., Kozak, C. A., and Nebert, D. W. (1990). Decreased *Hnf-1* gene expression in mice homozygous for a 1.2-centiMorgan deletion on chromosome 7. *DNA Cell Biol.* **9:** 771–776.
5. Green, E. L. (1981). "Genetics and Probability in Animal Breeding Experiments," pp. 77–113, MacMillan, New York.
6. Hirai, Y., Takebe, K., Takashina, M., Kobayashi, S., and Takeichi, M. (1992). Epimorphin: A mesenchymal protein essential for epithelial morphogenesis. *Cell* **69:** 471–481.
7. Hirai, Y. (1993). Molecular cloning of human epimorphin: Identification of isoforms and their unique properties. *Biochem. Biophys. Res. Commun.* **191:** 1332–1337.
8. Klinga-Levan, K., Levan, G., Stahl, F., Szpirer, J., and Szpirer, C. (1995). Rat gene map 1994. *Rat Genome* **1:** 15–47.
9. Kozak, C. A., Peyser, M., Krall, M., Mariano, T. M., Kumar, C. S., Pestka, S., and Mock, B. A. (1990). Molecular genetic markers spanning mouse chromosome 10. *Genomics* **8:** 519–524.
10. Pelham, H. R. B. (1993). Is epimorphin involved in vesicular transport? *Cell* **73:** 425–426.
11. Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., and Gray, J. (1988). Fluorescence *in situ* hybridization with human chromosome-specific libraries: Detection of trisomy 21 and translocations of chromosome 4. *Proc. Natl. Acad. Sci. USA* **85:** 9138–9142.
12. Stephanova, E., Tissir, F., Dusetti, N., Ivonna, J., Szpirer, J., and Szpirer, C. (1996). The rat genes encoding the pancreatitis-associated proteins I, II and III (*Pap1, Pap2, Pap3*) and the lithostathin/pancreatic stone protein/regeneration protein (*Reg*) colocalize at 4q33–q34. *Cytogenet. Cell Genet.* **72:** 83–85.
13. Szpirer, J., Levan, G., Thorn, M., and Szpirer, C. (1984). Gene mapping in the rat by mouse-rat somatic cell hybridization: Synteny of the albumin and alpha-fetoprotein genes and assignment to chromosome 14. *Cytogenet. Cell Genet.* **38:** 142–149.