J. Biochem. 120, 797-802 (1996)

Structural Organization of the *rae28* Gene, a Putative Murine Homologue of the *Drosophila polyhomeotic* Gene[1]

Md. Abdul Motaleb,* Yoshihiro Takihara,* Midori Nomura,* Yoichi Matsuda,† Toru Higashinakagawa,‡ and Kazunori Shimada*,[2]

*Department of Medical Genetics, Division of Molecular Biomedicine, Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita 565; †Laboratory of Animal Genetics, School of Agricultural Sciences, Nagoya University, Chikusa-ku, Nagoya 464-01; and ‡Mitsubishi Kasei Institute of Life Sciences, Machida, Tokyo 194

Received for publication, June 3, 1996

A putative murine homologue of the *Drosophila polyhomeotic* gene, named *rae28*, has been isolated from a genomic library of 129/SV mouse and its structural organization has been analyzed. *rae28* is a single gene of approximately 22 kb long and consists of 15 exons. Its 5′-flanking region lacks typical transcriptional regulatory sequences, such as TATA and CCAAT boxes, but contains GC-rich sequences and seven putative binding sites for a transcription factor, Sp1. One major transcription start point has been determined. The overall exon-intron organization suggested that three different Rae28 mRNAs are generated through alternative splicing. Furthermore, the *rae28* gene has been located on the R-positive F3 band of mouse chromosome 6 by the direct R-banding fluorescence *in situ* hybridization methods.

Key words: embryogenesis, F9 cell, genomic organization, nucleotide sequence, *polycomb* group genes.

In *Drosophila*, the *polycomb* group (*Pc*-G) genes are required for maintenance of correct spatial and temporal expression of *homeotic* genes during development, presumably through regulation of the high order chromatin structure (*1, 2*). The molecular mechanisms regulating embryogenesis and morphogenesis have been suggested to have been conserved from *Drosophila* to mammals (*3, 4*). The mammalian homologues of *Drosophila Pc*-G genes hitherto identified are the *bmi-1* (*5*), *mel-18* (*6*), and *M33* genes (*7*). Recently, we reported that one of the retinoic acid (RA)-inducible cDNAs isolated from mouse embryonal carcinoma F9 cells, named Rae28, encodes a novel protein sharing several characteristic motifs and highly homologous regions with a *Drosophila polyhomeotic* protein (*8, 9*). Based on these structural features, the *rae28* gene was presumed to be a murine homologue of the *Drosophila polyhomeotic* gene (*9*). Studies on mammalian homologues of the *Drosophila Pc*-G genes should provide clues for elucidating the molecular mechanisms underlying mammalian morphogenesis and embryogenesis. As the first of studies on the roles of the *rae28* gene in mouse development, we analyzed its structural organization in this work.

MATERIALS AND METHODS

Screening of Genomic Clones—A mouse 129/SV genomic library constructed from liver DNAs in the λFIX II vector (Stratagene) was screened by plaque hybridization under stringent conditions, using ^{32}P-labeled Rae2819 cDNA (*9*) as a probe. The positive phage clones were purified, and analyzed by restriction mapping and sequencing.

DNA Sequencing Analysis—The genomic DNA fragments that hybridized to the Rae2819 cDNA probe were subcloned into a pBluescript II KS plasmid (Stratagene), and then their sequences were determined by the dideoxy chain termination method (*10, 11*) and the cycle sequence method, using a 373A automated DNA sequencer (Perkin Elmer, Applied Biosystems Division).

Cells—Mouse embryonal carcinoma F9 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 15% fetal calf serum, and were induced to differentiate by treatment with 10^{-6} M RA (Sigma) (*12*).

Primer Extension Analysis—The 16 mer complementary to the sequence of the 3′ portion of exon 1 from nucleotide +114 to +129, 5′-CTTCGCGCCCGGCTCC-3′ (see Fig. 4B), was end-labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase. The labeled oligonucleotide was annealed to total cellular RNA (25 μg) extracted from F9 cells treated with RA for 3 h by the acid guanidium thiocyanate-phenol-chloroform method (*13*). The extension reaction

[1] This work was supported by Grants-in-Aid for Scientific Research (04454170 and 06807014), and a Grant-in-Aid for Creative Basic Research from the Ministry of Education, Science, Sports and Culture of Japan. This work was also supported by a Grant-in-Aid from the Japan Medical Association and one from the Uehara Memorial Foundation. All the nucleotide sequences reported herein have been deposited in the DDBJ DNA database under accession No. D78515.

[2] To whom correspondence should be addressed. Phone: +81-6-879-8324, Fax: +81-6-879-8326

Abbreviations: ESTs, expressed sequence tags; FISH, fluorescence *in situ* hybridization; *Pc*-G, *polycomb* group; DMEM, Dulbecco's modified Eagle's medium; RA, retinoic acid; RAREs, retinoic acid-responsive elements; Rae28, retinoic acid early-inducible cDNA clone 28; *rae28*, gene corresponding to Rae28 cDNA; RAE28, protein coded by *rae28*; 1×SSC, 0.15 M NaCl-0.015 M sodium citrate, pH 7.6; *tsp*, transcription start point(s).

was carried out with avian myeloblastosis virus reverse transcriptase (Life Science) at 42°C for 60 min. The extended product was analyzed by running on a 6% sequencing gel alongside a dideoxy sequencing ladder generated using the same primer and a genomic DNA fragment subcloned into a pBluescript II KS plasmid.

Zoo Blot Analysis—High molecular weight DNAs were extracted from human peripheral polymorphocytes, mouse liver, chicken embryos, *Xenopus* liver, *Drosophila* Schneider cells, and yeast (*14*). These DNAs were digested with *Eco*RI, separated on a 0.8% agarose gel and then transferred to a Nylon membrane, Hybond-N (Amersham). The blots were hybridized with the ^{32}P-labeled 3′-side probe of Rae2819 cDNA at 55°C and subsequently washed with 2× SSC at the same temperature. Bands were visualized by autoradiography.

Fluorescence In Situ Hybridization (FISH) Method—The R-banded chromosome slides were denatured and dehydrated in a 70-85-100% ethanol series. The 16-kb genomic DNA fragment, derived from clone R5 (Fig. 1B), was labeled by nick translation with biotin 16-dUTP (Boehringer Mannheim). The probe was hybridized on the denatured slides overnight at 37°C. After washing, the slides were incubated under coverslips with fluoresceinated avidin at 1:500 dilution in 1% BSA/4×SSC and then stained with 0.75 μg/ml propidium iodide. Excitation at wavelengths of 450-490 nm (Nikon filter set B-2A) and near 365 nm (UV-2A) was performed for observation of the R-banding and G-banding patterns, respectively (*15, 16*).

	Exon		Intron		Exon	
		3′ 5′		3′ 5′		
1	GGGCGCGAAG	GTAACCGGGC	-----//-----	TTCTGTCTAG	GGCTTGAGT	2
2	AGCAGTGCAG	GTGAGGCTCA	-----//-----	TGTCTTCTAG	GCTCTGCAA	3
3	CAGGTGAGAG	GTCAGCAGCG	-----//-----	CCTTGAACAG	GCCACGATT	4
(3a	TGTCCAGCAG	GTGAGAGGTC)				
4	CCCAGGCCTC	GTGAGTACTG	-----//-----	TACTGTGCAG	ATGAATCTC	5
5	TGTATCTGCG	GTAAGTCACA	-----//-----	CTCCTCGCAG	GCCACAGCT	6
6	TGCAGATCAG	GTCAGTGGCA	-----//-----	TATCTTCCAG	GTGCAGAAC	7
7	ATTAGCTCAG	GTAAGGTGTC	-----//-----	TCACACATAG	CCACATACA	8
8	CTGCCGGGTA	GTATGCTAGT	-----//-----	TTCCTCTTAG	GAAAAGCCC	9
9	AGTCTTGGAG	GTGAGTAGCC	-----//-----	TGTTTTTCAG	AGAAAGCTG	10
10	GCCTTTTCCG	GTGAGGGCAG	-----//-----	TCTGGAGCAG	GTGGGATGT	11
11	CCCTCTGTAG	GTAAGTGATG	-----//-----	GTTTTCCAAG	TTAGAGAAG	12
12	GTGCAAAGAG	GTACTGCGCC	-----//-----	TCATCTGTAG	GTACAATGT	13
13	CGCCATCGGG	GTGAGCTGCT	-----//-----	TCCTCTCAAG	CAAGAGGAC	14
14	CTTCTCTGCA	GTACGAGGTG	-----//-----	TTCCTTTTAG	GCTGCCAAG	15

Fig. 2. **Nucleotide sequences of the predicted splice junctions of the *rae28* gene.** The locations of exon and intron sequences are indicated. 5′ and 3′ indicate the predicted termini of each exon and intron, and the numbers on either side of the figure indicate the exon numbers. In parenthesis is the nucleotide sequence at the alternative splice junction predicted at the 3′ boundary of exon 3. The nucleotide sequences are underlined at the splice junctions of the 5′ and 3′ boundaries of intron 9, where alternative splicing was predicted.

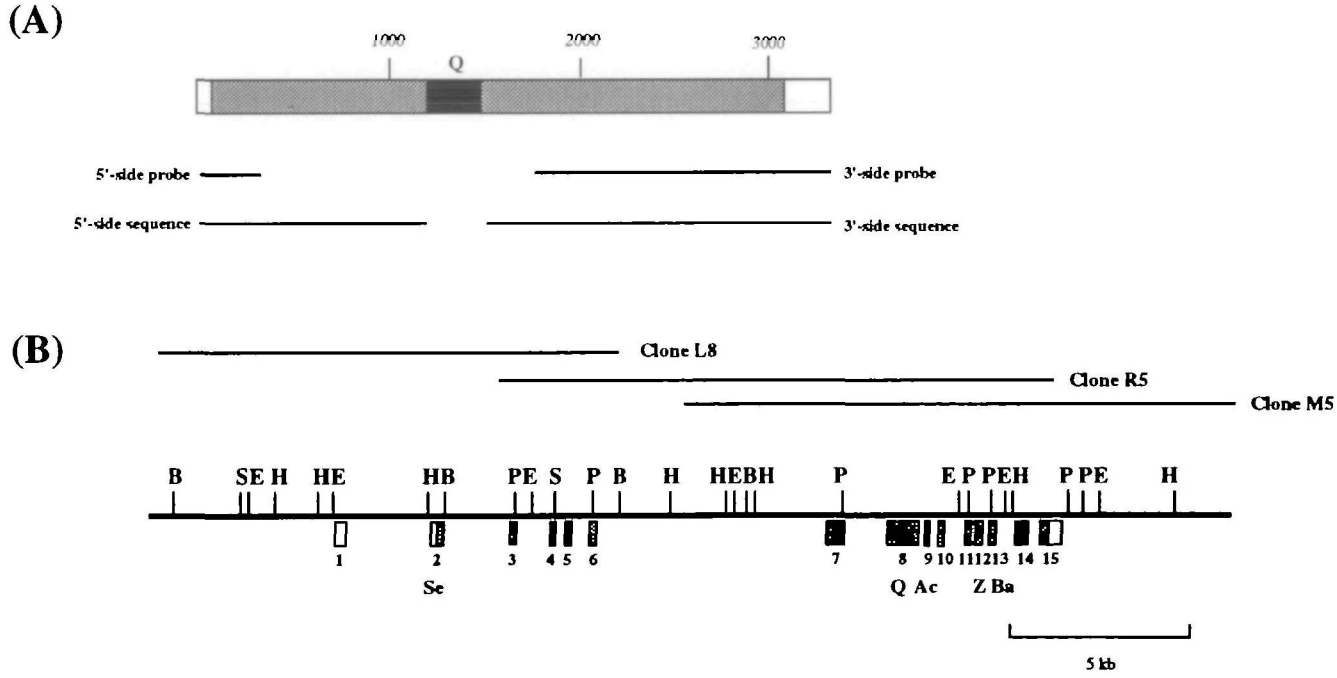

Fig. 1. **Structural organization of the *rae28* gene.** (A) Schematic presentation of Rae2819 cDNA. The numbers indicate the nucleotide numbers of Rae2819 cDNA (*9*). The 5′-side probe was a 0.3 kb *Eco*RI-*Sal*I fragment and the 3′-side probe a 1.5 kb *Pvu*II-*Eco*RI fragment (*9*). Two sequences were used for the database search: one was a 5′-side sequence covering nucleotide +1 to +1191, and the other a 3′-side sequence covering nucleotide +1491 to +3290. The symbols are as follows: open boxes, untranslated regions; shaded boxes, coding regions; solid box, sequence encoding the glutamine-rich region. (B) Genomic DNAs carried by three λ clones, restriction map, and complete exon/intron organization of the *rae28* gene. The numbers represent exon numbers. Restriction sites: E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII; S, *Sal*I; P, *Pst*I. The symbols are as follows: Se, serine-rich region; Q, glutamine-rich region; Ac, acidic amino acid-rich region; Z, single zinc finger; Ba, basic amino acid-rich region. The other symbols are the same as in (A).

RESULTS AND DISCUSSION

Structure and Sequence of the rae28 Gene—Rae2819 cDNA, one of the full-length Rae28 cDNAs, contains sequences encoding a glutamine-rich region (Fig. 1A) (*9*). Southern blot analysis of mouse genomic DNAs with this cDNA as a probe showed smear bands (data not shown), probably due to the cross hybridization with various genes encoding proteins with glutamine-rich regions (*17*). So, we prepared the following two DNA fragments, which lack the sequences encoding the glutamine-rich region, and used them as probes: one covers the 5′- side, and the other the 3′-side (Fig. 1A) (*9*). A genomic library of 129/SV mouse was screened using these two probes and a series of λ phage clones was isolated. Three of the isolates overlapped and comprised a genomic DNA covering the entire *rae28* gene (Fig. 1B). The inserts of the clones hybridized with the cDNA probes were subcloned into pBluescript II KS, and then subjected to further fine restriction mapping and sequence analysis (*14*). The restriction map and alignment of the cloned DNAs are shown in Fig. 1B. The exon-intron junctions were determined precisely by comparing the genomic sequences with that of the Rae2819 cDNA, because it corresponds to one of the major transcripts of the *rae28* gene (*9*). The overall exon-intron organization of the *rae28* gene is schematically shown in Fig. 1B. It consists of 15 exons and spans approximately 22 kb. All the major bands detected on Southern blot analysis (*9*) were compatible with this restriction map, indicating that *rae28* is a single gene. The genomic organization of the *rae28* gene is different from that of the *Drosophila polyhomeotic* gene, because in *Drosophila*, two duplicated highly homologous *polyhomeotic* genes are arranged as a tandem repeat in a single chromosomal locus (*18*).

The first ATG initiation codon is located 45 bp downstream from the 5′ end of exon 2, and the surrounding sequences fulfill Kozak's rule (data not shown) (*19*). This codon corresponds to the putative translation initiation codon found in Rae2819 cDNA (*9*). The deduced RAE2819 protein contains five characteristic motifs (*9*). Interestingly, each these motif is encoded by a different exon, *i.e.*, the serine-rich region is encoded by exon 2, the glutamine-rich region by exon 8, the acidic amino acid-rich region by exon 9, the single zinc finger by exon 12, and the basic amino acid-rich region by exon 13 (Fig. 1B).

The nucleotide sequences of all the exon-intron junctions fulfill the GT-AG rule (Fig. 2) (*20*). We previously isolated three different Rae28 cDNAs, *i.e.*, Rae2803, Rae2819, and Rae2821 cDNAs (*9*). The structure of the *rae28* gene suggested that the Rae2803 and Rae2821 cDNAs correspond to mRNAs generated through alternative splicing (Fig. 3A): Rae2803 cDNA corresponds to mRNA generated through alternative splicing generating a seven-nucleotide insert between exons 3 and 4, and deletion of exon 6 (Fig. 3B); Rae2821 cDNA corresponds to mRNA which retains

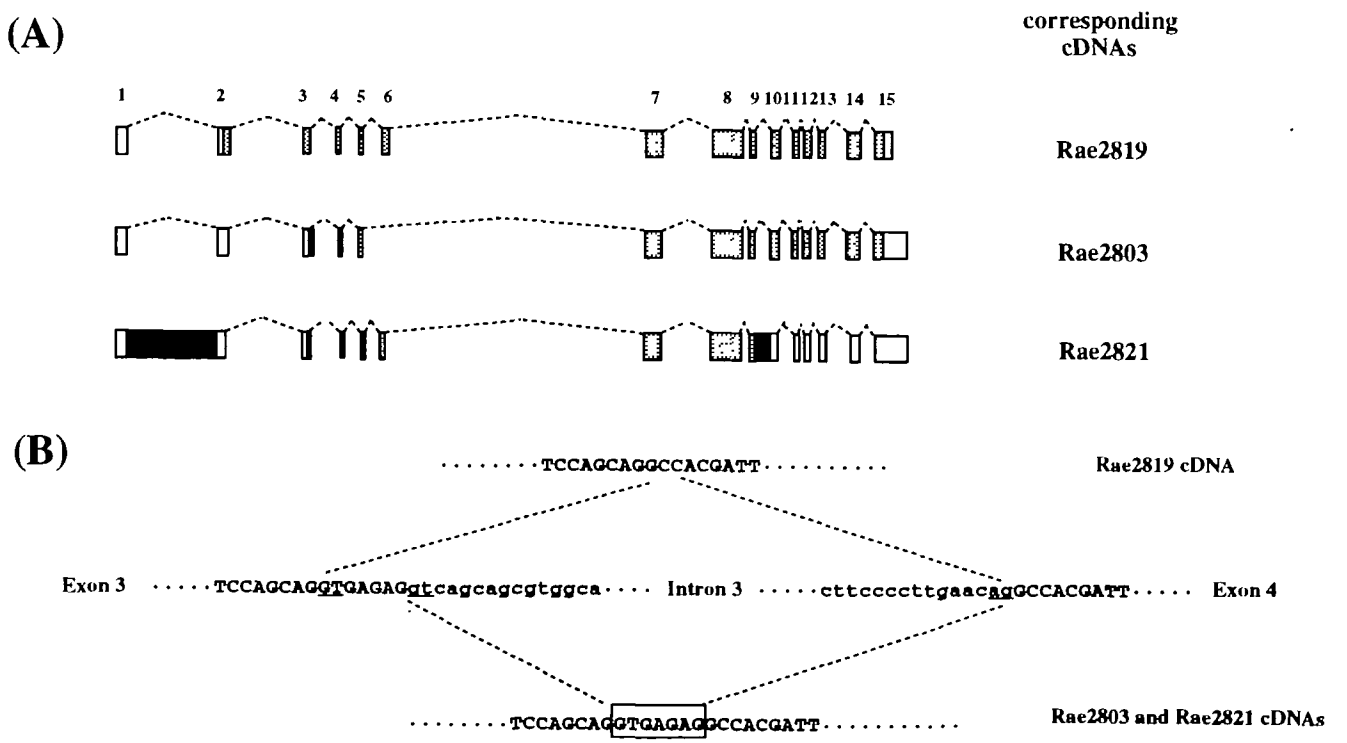

Fig. 3. **Alternative splicing of the *rae28* gene transcripts.** (A) Three different splicing patterns were proposed to explain the generation of the Rae2819, Rae2803, and Rae2821 cDNAs. The symbols are as follows: open boxes, untranslated regions; dotted boxes, coding regions; solid boxes, exons generated through alternative splicing. The numbers represent exon numbers. (B) The nucleotide sequences of the exon-intron junctions of intron 3 and the alternative splicing predicted in the 3′ region of exon 3. The nucleotides in exon and intron regions are represented by capital and lower-case letters, respectively. The conserved nucleotide sequences in the splicing sites are indicated by underlines. The predicted splicing is shown by dotted lines, and the seven-nucleotide insert between exons 3 and 4 is enclosed by an open box.

intron 1, the same seven-nucleotide insert and intron 9 (Fig. 3B). These two cDNAs presumably encode truncated forms of the RAE28 protein, due to the existence of stop codon(s) in their open reading frames.

Exon 15 contains two putative polyadenylation signals: one, AATAAA, at 172 bp and the other, ATTAAA, at 717 bp downstream from the stop codon (data not shown) (*21*).

Characterization of the 5′-Flanking Region—To deter-

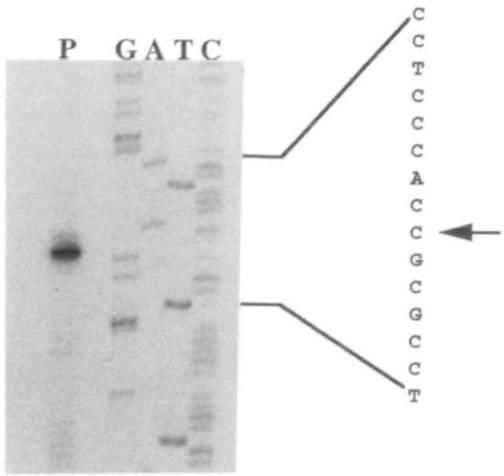

Fig. 4. **Identification of *tsp* and the nucleotide sequence of the 5′-flanking region.** (A) Primer extension analysis of the 5′-end region of the *rae28* gene. The extension products are presented with the sequence ladders. Lanes G, A, T, and C are sequence ladders, and lane P, primer extension products. The arrow on the right indicates the major extension product. (B) The nucleotide sequence of the 5′-flanking region. The major *tsp* is indicated by a bent arrow. The exon sequence is represented by capital letters. The potential binding sites for the transcription factors (AP2, Ets-1, Sp1, *etc.*) are underlined. The boxed sequence was used as a primer for primer extension analysis. The nucleotide sequence is numbered relative to the major *tsp* at +1. The nucleotides upstream from the *tsp* are negatively numbered.

mine the transcription start point(s) (*tsp*), we performed primer extension analysis (Fig. 4A) (*22*). The primer shown in Fig. 4B was labeled at its 5′ end, hybridized with total cellular RNA extracted from F9 cells treated with RA for 3 h, and then extended with reverse transcriptase (*9*). One strong and several faint bands were detected, the strong band corresponding to the mRNA starting from nucleotide −130 (Fig. 4A).

Sequencing analysis revealed that the 5′-flanking region contains neither a TATA box nor a CCAAT box. It contains GC-rich regions and apparently contains a large number of CpG dinucleotides: 86% of the nucleotides from position −1 to −300 are GC (Fig. 4B). The 5′-flanking region contains putative binding sites for the following transcription factors: seven Sp1 sites (*23*), six AP2 sites (*24*), two Ets-1 sites (*25*), two Myb sites (*26*), two GAGA sites (*27*), and one CREB site (*28*) (Fig. 4B). Although transcription of the *rae28* gene is transiently induced during RA-mediated F9 cell differentiation (*9*), we could not detect any sequences homologous to RA-responsive elements (RAREs) in its 5′-flanking region, from nucleotide −1 to −1389 (*29*).

Zoo Blot Analysis of the rae28 Gene—Since the deduced RAE28 protein shared several characteristic structures with a *Drosophila polyhomeotic* protein (*9*), we examined whether or not the *rae28* gene has been conserved during evolution. High molecular weight DNAs were prepared from man, mouse, chicken, *Xenopus*, *Drosophila*, and yeast. These DNAs were subjected to zoo blot analysis under low stringency conditions, using the 3′-side probe of Rae2819 cDNA (Fig. 1A). As shown in Fig. 5, we detected one to several discrete bands, not only in mouse, but also in man, chicken and *Drosophila* DNAs. These results suggest that the *rae28* gene has been conserved during evolution.

Human Expressed Sequence Tags (ESTs) Showing Similarities to Rae2819 cDNA—We performed a database

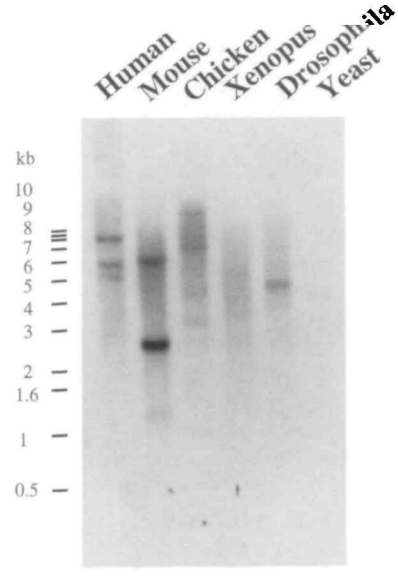

Fig. 5. **Zoo blot analysis of the *rae28* gene.** The species of each DNA is shown at the top of the lane and the lengths of size markers at the left.

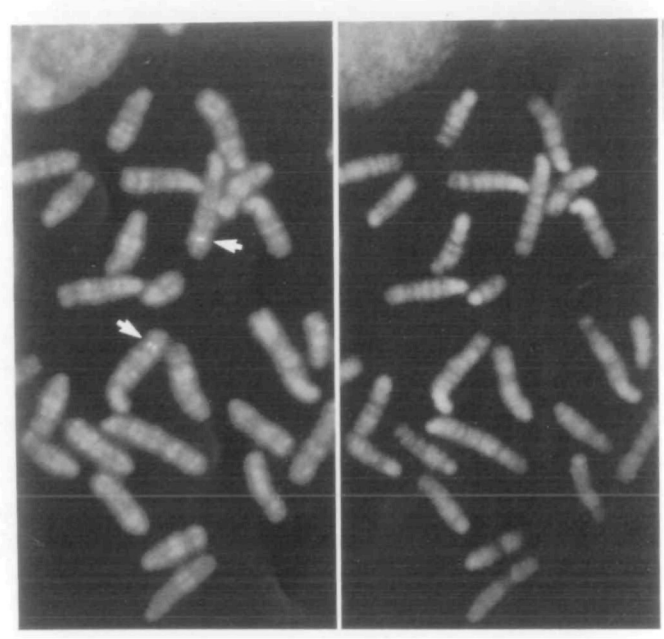

Fig. 6. **Chromosomal localization of the *rae28* gene on R-banded chromosomes.** Left panel: a representative R-banded metaphase spread after *in situ* hybridization with a *rae28* genomic probe. The arrows indicate hybridization signals of both alleles on chromosome 6. Right panel: a G-banded diagram of the same metaphase spread.

TABLE I. Human ESTs showing similarities to Rae2819 cDNA.

GenBank AC #[a]	Source of cDNA library[b]	Blastn scores[c]	p[d]	Homologous regions of Rae2819 cDNA[e]	ID (%[f])
R15428	Infant brain	1094	$3.1e^{-83}$	2120-2399	87
T52131	Fetal spleen	854	$2.9e^{-73}$	2183-2437, 2494-2520	
				2527-2546	83
N41621	Placenta, 8 to 9 weeks	582	$7.9e^{-70}$	2929-3056, 3058-3100	
				3109-3161, 3209-3289	85
T09455	Infant brain	534	$1.1e^{-34}$	2878-3100	70
R12754	Infant brain	152	$4.4e^{-13}$	3067-3100, 3109-3161	
				3209-3256	79
H69928	Olfactory epithelium	246	$4.0e^{-09}$	3209-3289	76

[a]Sequences showing similarities to Rae2819 cDNA were searched for in the non-redundant nucleic acid sequence database and the database of ESTs (= DBEST), using the BLAST program mail server (*30*) at either the National Center for Biotechnology Information, USA, or the Human Genome Center, Institute of Medical Science, The University of Tokyo, Japan, and the GenBank accession numbers of the matched human ESTs are shown. These searches were made from February 1-3, 1996. [b]The sources of mRNAs used to prepare the cDNA libraries are shown. [c]Blastn scores and [d]Poisson p-values are shown. The p-value, <0.01, is considered statistically significant. [e]The nucleotide positions of the Rae2819 cDNA showing similarities to each human EST are indicated. [f]The percentage of identical nucleotides per total nucleotides present within the homologous region is shown.

search using the BLAST program (*30*) and the following two parts of the Rae2819 cDNA sequence: one covering the 5′-side sequence, and the other the 3′-side sequence (Fig. 1A). Six human ESTs showed significant similarities to the 3′-side sequence, suggesting that they are human counterparts of the Rae28 cDNA (Table I). Moreover, these results suggest that the human counterpart(s) of the *rae28* gene is transcribed in infant brain, fetal spleen, placenta, and olfactory epithelium. A similar search using the 5′-side sequence did not reveal any homologous sequences.

Chromosomal Mapping of the rae28 Locus—Chromosomal assignment of the mouse *rae28* gene was carried out by the R-banding fluorescence *in situ* hybridization (FISH) method, using a 16-kb genomic DNA fragment prepared from clone R5 as a probe (Fig. 1B). As shown in Fig. 6, the signals were localized to the R-positive F3 band of chromosome 6, which indicated that the *rae28* gene consists of a single genetic locus (*15*). This locus is syntenic to human chromosome 12p11-p13 (*31*).

In conclusion, the *rae28* gene is a conserved single gene consisting of 15 exons, which is located in the F3 region of chromosome 6. Three different mRNAs may be generated through alternative splicing. Although expression of the *rae28* gene is inducible with RA (*9*), RAREs were not found in its 5′ flanking region. It remains to be elucidated whether the expression of the *rae28* gene is regulated by RA indirectly or there are RAREs in a region other than the 5′-flanking region examined.

At present there is no mutant of the *rae28* locus. To elucidate the physiological function of the gene, examination of *rae28*-deficient mice, generated by targeted mutagenesis, is now in progress in our laboratory.

We wish to thank Dr. T. Yasunaga and K. Sumitomo for their help in the computer analyses, and M. Ohara for the helpful comments. We also thank the National Center for Biotechnology Information, USA, and the Human Genome Center, Institute of Medical Science, The University of Tokyo, Japan, for providing access to the network BLAST server.

REFERENCES

1. Struhl, G. and Akam, M. (1985) Altered distributions of *Ultrabithorax* transcripts in *extra sex combs* mutant embryos of *Drosophila*. *EMBO J.* **4**, 3259-3264
2. Paro, R. (1993) Mechanisms of *heritable* gene repression during development of *Drosophila*. *Curr. Opin. Cell Biol.* **5**, 999-1005
3. Kenyon, C. (1994) If birds can fly, why can't we? *Homeotic* genes and evolution. *Cell* **78**, 175-180
4. Krumlauf, R. (1994) *Hox* genes in vertebrate development. *Cell* **78**, 191-201
5. Brunk, B.P., Martin, E.C., and Adler, P.N. (1991) *Drosophila* genes *Posterior Sex Combs* and *Suppressor two of zeste* encode proteins with homology to the murine *bmi-1* oncogene. *Nature* **353**, 351-353
6. Tagawa, M., Sakamoto, T., Shigemoto, K., Matsubara, H., Tamura, Y., Ito, T., Nakamura, I., Okitsu, A., Imai, K., and Taniguchi, M. (1990) Expression of novel DNA-binding protein with zinc finger structure in various tumor cells. *J. Biol. Chem.* **265**, 20021-20026
7. Pearce, J.J.H., Singh, P.B., and Gaunt, S.J. (1992) The mouse has a *Polycomb*-like *chromobox* gene. *Development* **114**, 921-929
8. Dura, J.-M., Randsholt, N.B., Deatrick, J., Erk, I., Santamaria, P., Freeman, J.D., Freeman, S.J., Weddell, D., and Brock, H.W. (1987) A complex genetic locus, *polyhomeotic*, is required for segmental specification and epidermal development in *D. melanogaster*. *Cell* **51**, 829-839
9. Nomura, M., Takihara, Y., and Shimada, K. (1994) Isolation and characterization of retinoic acid-inducible cDNA clones in F9 cells: One of the early inducible clones encodes a novel protein sharing several highly homologous regions with a *Drosophila Polyhomeotic* protein. *Differentiation* **57**, 39-50
10. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
11. Hattori, M. and Sakaki, Y. (1986) Dideoxy sequencing method using denatured plasmid templates. *Anal. Biochem.* **152**, 232-238
12. Rudnicki, M.A. and McBurney, M.W. (1987) Cell culture methods and induction of differentiation of embryonal carcinoma cell lines in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (Robertson, E.J., ed.) pp. 19-49, IRL Press, Oxford
13. Chomczynski, P. and Sacchi, N. (1987) Single step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156-159
14. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
15. Matsuda, Y., Harada, Y.-N., Natsuume-Sakai, S., Lee, K., Shiomi, T., and Chapman, V.M. (1992) Location of the mouse *complement factor H* gene (*cfh*) by FISH analysis and replication R-banding. *Cytogenet. Cell Genet.* **61**, 282-285
16. Matsuda, Y. and Chapman, V.M. (1995) Application of fluorescence *in situ* hybridization in genome analysis of the mouse. *Electrophoresis* **16**, 261-272
17. Kadonaga, J.T., Carner, K.R., Masiarz, F.R., and Tjian, R. (1987) Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51**, 1079-1090
18. Deatrick, J., Daly, M., Randsholt, N.B., and Brock, H.W. (1991) The complex genetic locus *polyhomeotic* in *Drosophila melanogaster* potentially encodes two homologous zinc-finger proteins. *Gene* **105**, 185-195
19. Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**, 857-872
20. Sharp, P.A. (1981) Speculations on RNA splicing. *Cell* **23**, 643-646
21. Wahle, E. and Keller, W. (1992) The biochemistry of 3′-end cleavage and polyadenylation of messenger RNA precursors. *Annu. Rev. Biochem.* **61**, 419-440
22. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1987) *Current Protocols in Molecular Biology*, section 4.8, Greene Publishing Associates and Wiley-Interscience, New York
23. Pugh, B.F. and Tjian, R. (1991) Transcription from a TATA-less promoter requires a multisubunit *TFIID* complex. *Genes Dev.* **5**, 1935-1945
24. Haslinger, A. and Karin, M. (1985) Upstream promoter element of the human *metallothionein*-IIA gene can act like an enhancer element. *Proc. Natl. Acad. Sci. USA* **82**, 8572-8576
25. Wasylyk, B., Wasylyk, C., Flores, P., Begue, A., Leprince, D., and Stehelin, D. (1990) The *c-ets* proto-oncogenes encode transcription factors that cooperate with *c-Fos* and *c-Jun* for transcriptional activation. *Nature* **346**, 191-193
26. Ness, S.A., Marknell, A., and Graf, T. (1989) The *v-myb* oncogene product binds to and activates the promyelocyte-specific *mim-1* gene. *Cell* **59**, 1115-1125
27. Biggin, M.D. and Tjian, R. (1988) Transcription factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. *Cell* **53**, 699-711
28. Sassone-Corsi, P. (1988) Cyclic AMP induction of early adenovirus promoters involves sequences required for E1A trans-activation. *Proc. Natl. Acad. Sci. USA* **85**, 7192-7196
29. Leid, M., Kastner, P., and Chambon, P. (1992) Multiplicity generates diversity in the retinoic acid signaling pathways. *Trends Biochem. Sci.* **17**, 425-433
30. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410
31. Elliott, R.W. and Moore, K.J. (1994) Mouse chromosome 6. *Mamm. Genome* **5**, S79-S103