Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 3594–3598, May 1990
Biochemistry

Isolation and characterization of the human homologue of *rig* and its pseudogenes: The functional gene has features characteristic of housekeeping genes

(insulinoma gene/CpG islands)

KIYOTO SHIGA, HIROSHI YAMAMOTO, AND HIROSHI OKAMOTO*

Department of Biochemistry, Tohoku University School of Medicine, Sendai 980, Miyagi, Japan

Communicated by Sidney Weinhouse, February 1, 1990 (received for review July 28, 1989)

ABSTRACT *rig* (*r*at *i*nsulinoma *g*ene) was first isolated from a cDNA library of rat insulinomas and has been found to be activated in various human tumors such as insulinomas, esophageal cancers, and colon cancers. Here we isolated the human homologue of *rig* from a genomic DNA library constructed from a human esophageal carcinoma and determined its complete nucleotide sequence. The gene is composed of about 3000 nucleotides and divided into four exons separated by three introns: exon 3 encodes the nuclear location signal and the DNA-binding domain of the RIG protein. The transcription initiation site was located at −46 base pairs upstream from the first ATG codon. The 5′-flanking region of the gene has no apparent TATA-box or CAAT-box sequence. However, two GC boxes are found at −189 and −30 base pairs upstream from the transcription initiation site and five GC boxes are also found in introns 1 and 2. The gene is bounded in the 5′ region by CpG islands, regions of DNA with a high GC content and a high frequency of CpG dinucleotides relative to the bulk genome. Furthermore, the human genome contains at least six copies of *RIG* pseudogenes, and four of them have the characteristics of processed pseudogenes. From these results together with the finding that *RIG* is expressed in a wide variety of tissues and cells, we speculate that *RIG* belongs to the class of "housekeeping" genes, whose products are necessary for the growth of all cell types.

rig (*r*at *i*nsulinoma *g*ene) has been isolated from a cDNA library of rat insulinomas (1) induced by the combined administration of streptozotocin and a poly(ADP-ribose) synthetase inhibitor (2, 3). *rig* encodes a 145-amino acid protein that contains a nuclear location signal and a DNA-binding domain (1, 4). *rig* was highly expressed not only in the chemically induced rat insulinomas but also in a BK virus-induced hamster insulinoma and spontaneously occurring human insulinomas (4, 5). The 145-amino acid sequence deduced from the cDNA sequences remained invariant in rat, hamster, and human proteins, suggesting that *rig* has evolved under extraordinarily strong selective constraints (4). Microinjection of antisense oligonucleotides complementary to *rig* mRNA suppressed DNA synthesis in hamster insulinoma cells (4). *RIG* was also activated in various human tumors such as esophageal cancers and colon cancers (6). Moreover, we have found that *rig* is transiently expressed in rat regenerating liver and in synchronously cultured rat hepatocytes: *rig* expression was stimulated prior to DNA synthesis and RIG protein accumulated in the nucleus during the S phase of the cell cycle (7). These results suggest that *rig* can be involved in a more general way in cell growth or replication (6, 8).

In the present study, we have isolated genomic *RIG* from a DNA library derived from a human esophageal carcinoma and determined the complete nucleotide sequence.† Available data suggest that *RIG* belongs to the class of "housekeeping" genes, whose products are necessary for the growth of all cell types.

MATERIALS AND METHODS

Materials. A tumor-containing portion of the esophagus was obtained from a patient (52-year-old male) at Tohoku University Hospital (Sendai) under an approved protocol. The tumor, diagnosed as squamous cell carcinoma of the esophagus, and the normal esophageal mucosa were excised from the surgical specimen and kept frozen below −80°C until the isolation of nucleic acids. Deoxycytidine 5′-[α-^{32}P]triphosphate ([α-^{32}P]dCTP) (3000 Ci/mmol; 1 Ci = 37 GBq), cytidine 5′-[α-^{32}P]triphosphate ([α-^{32}P]CTP) (800 Ci/mmol), and a Multiprime DNA labeling system were purchased from Amersham. λEMBL3, λEMBL4, λgt10, pBS, and T7 RNA polymerase were from Stratagene. Restriction enzymes, T4 DNA ligase, Klenow fragment of DNA polymerase I, and reverse transcriptase were from Takara Shuzo (Kyoto).

RNA Extraction and Northern Blot Analysis. Total cellular RNA was extracted as described (9), electrophoresed on a 1.5% (wt/vol) agarose gel containing 1.1 M formaldehyde, and transferred onto a nitrocellulose membrane filter. The filter was hybridized (10) to an RNA probe that had been synthesized (11) from human *RIG* cDNA insert in pBS (4) in the antisense orientation with T7 RNA polymerase and [α-^{32}P]CTP.

DNA Extraction and Southern Blot Analysis. High molecular weight DNA was extracted from tissues as described (12) and digested with restriction endonucleases. The digested DNA was electrophoresed on a 0.8% (wt/vol) agarose gel and transferred onto nitrocellulose filters (13, 14). Filters were hybridized (15) to a cDNA probe that had been labeled in the presence of [α-^{32}P]dCTP by the random priming method (16) to a specific activity of 1–2 × 10^9 cpm/μg. The cDNA probe was the 500-base pair (bp) *Eco*RI fragment of human *RIG* cDNA that had been isolated from a human pancreas cDNA library (17): the *Eco*RI fragment contains the entire coding region of the cDNA.

Isolation of Genomic Clones. Three different genomic libraries were constructed from the esophageal carcinoma DNA. The first library was constructed as described (18) with the partially *Sau*3AI-digested DNA, which was ligated to bacteriophage λEMBL3 DNA, packaged *in vitro*, and propagated in P2392, a strain of *Escherichia coli* K-12. The second

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.

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M32405).

library was constructed with the completely *Dra* I-digested, size-fractionated DNA [9–23 kilobase pairs (kbp)] and λEMBL4 in P2392, and the third library was constructed with the *Dra* I-digested, size-fractionated DNA (5.5–7.5 kbp) and λgt10 in *E. coli* Y1089. Libraries were screened as described (19) with the ^{32}P-labeled 500-bp *Eco*RI fragment of human *RIG* cDNA. Positively hybridized clones were plaque-purified and DNA was isolated from the recombinant phages as described (20).

Nucleotide Sequencing. Cloned DNA was cleaved with various restriction endonucleases and subcloned into pBS vector. The nucleotide sequences of restriction fragments were determined by the dideoxy chain-termination method using deoxy-7-deazaguanosine triphosphate as a substrate (21, 22).

RESULTS

Northern and Southern Blot Analyses of a Human Esophageal Carcinoma. We analyzed RNAs from a surgically removed human esophageal carcinoma and from its normal counterpart (i.e., esophageal epithelium) from the same patient using human *RIG* cDNA as probe. As shown in Fig. 1*A*, an RNA species that hybridized to *RIG* cDNA was present at a much higher level in the carcinoma than in the normal counterpart. The RNA was 0.7 kb long, the same size as human *RIG* mRNA in insulinomas (4). In Southern blot analysis of genomic DNA from the carcinoma (Fig. 1*B*), *Bam*HI, *Bgl* II, *Dra* I, *Eco*RI, *Eco*RV, *Hin*cII, *Hpa* I, and *Xmn* I generated several restriction fragments that hybridized to human *RIG* cDNA. Southern blots of *Dra* I-digested DNA from the normal counterpart showed the same pattern of positive bands (14.0, 6.8, 4.0, 3.2, 1.8, 1.0, and 0.8 kbp) as that obtained with the carcinoma DNA (Fig. 1*C*), indicating that the gene had been neither amplified nor extensively rearranged in the carcinoma.

Isolation and Nucleotide Sequence Determination of Human *RIG*. A genomic DNA library constructed with 5.5- to 7.5-kbp fragments of *Dra* I-digested esophageal carcinoma DNA and λgt10 was screened with human *RIG* cDNA as probe. Six positively hybridized clones were isolated, and these clones were found to have an insert with the same restriction map and thought to correspond to the 6.8-kbp *Dra* I fragment shown in Fig. 1*C*. The DNA insert was excised from the vector DNA and subcloned. As shown in Fig. 2, all of the bases of the 2.8-kbp fragment (nucleotides −549 to 2313) were sequenced. Comparison of the genomic sequence with the human *RIG* cDNA sequence (4) revealed that the coding region is divided into four exons separated by three introns: the nucleotide sequence of the coding region of the gene is identical to that of the cDNA. As shown in Fig. 2*B*, exon 1 of the gene is 49 bp long and encodes the 5′-untranslated region of the *RIG* mRNA and the initiator methionine. Exon 2 is 86 bp long and encodes the amino-terminal portion of the RIG protein. Exon 3 is the longest exon at 235 bp long and encodes both the nuclear location signal and the DNA-binding domain of the RIG protein (4). Exon 4, 145 bp long, encodes the carboxyl-terminal portion of the protein and the 3′-untranslated region of the mRNA. Introns 1, 2, and 3 are 377, 1125, and 94 bp long, respectively. All exon–intron junctions conform to the "GT-AG" rule (24). *Alu* sequences were found at the 5′ end of the sequence (nucleotides −549 to −500) and in the middle of intron 2 (nucleotides 901–1310).

Features of the Promoter Region of Human *RIG*. A synthetic oligonucleotide was used in a primer extension reaction to define the 5′ end of the gene transcript. As shown in Fig. 3, a primer-extended product 35 nucleotides in length was obtained with RNA from a human esophageal carcinoma. S1 nuclease mapping was used to further define the end of *RIG* mRNA molecules and similar results were obtained (data not shown). Thus, we concluded that the cap site was at the C/G pair, 46 bp upstream from the first ATG codon (Fig. 2*B*). Neither apparent TATA-box nor CAAT-box sequences existed in the ≈540-bp 5′-flanking region, but two GC-box sequences were found 189 and 30 bp upstream from the cap site. GC-box sequences were also found in introns 1 and 2. Upstream from the transcription initiation site, there were several kinds of direct or inverted repeats: TTTTTGTA, at nucleotides −496 to −489, −455 to −448, −445 to −438; TTTTAT, at −486 to −481, −478 to −473, −391 to −386; TCCCAC.GTGGGA, at −311 to −299; GACTT. . . . . AAGTC, at −182 to −168; GGCGCG. . . . CGCGCC, at −81 to −66; GTGCGCGCAC, at −48 to −39. A heptamer sequence, GTAGGCG, found in the human hypoxanthine phosphoribosyltransferase gene promoter (25), was noted at position −265. The DNA sequence from the 5′-flanking region 201 bp upstream from the transcription initiation site to intron 2 (nucleotides −201 to 664) has an average G+C content of as high as 70% (Fig. 4). Furthermore, within this segment of the gene the ratio of observed to expected frequencies of the CpG dinucleotide was found to be extremely high, with the average value of 1.00 (Fig. 4) indicating that there has been

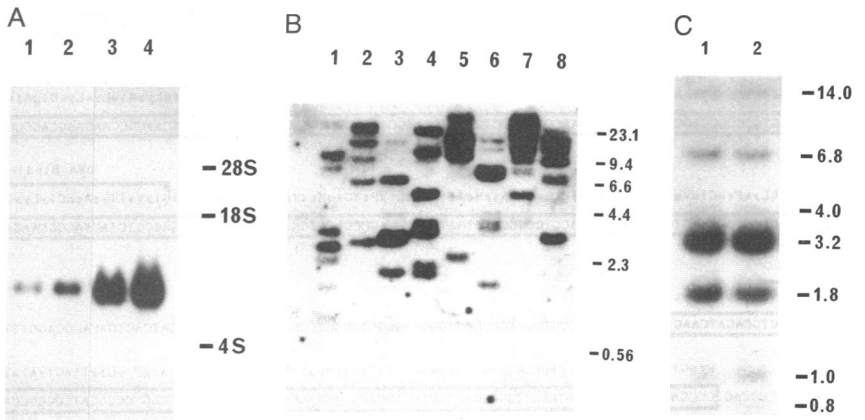

FIG. 1. (*A*) Northern blot analysis of *RIG* mRNA in human esophageal mucosa and carcinoma. Lanes 1 and 2, RNA from human esophageal mucosa (12.5 and 25 μg, respectively); lanes 3 and 4, RNA from human esophageal carcinoma (12.5 and 25 μg, respectively). Bars indicate 28S, 18S, and 4S RNAs on the same gel. (*B*) Southern blot analysis of human genomic *RIG*. Ten micrograms of DNA from human esophageal carcinoma was digested with *Bam*HI (lane 1), *Bgl* II (lane 2), *Dra* I (lane 3), *Eco*RI (lane 4), *Eco*RV (lane 5), *Hin*cII (lane 6), *Hpa* I (lane 7), and *Xmn* I (lane 8) and hybridized with human *RIG* cDNA. Molecular standard markers are indicated at the right in kbp. (*C*) Southern blot analysis of *Dra* I-digested DNA. Ten micrograms of DNA from human esophageal mucosa (lane 1) and carcinoma (lane 2) from the same patient was digested with *Dra* I and hybridized with human *RIG* cDNA. The size of each hybridizing band is indicated at the right in kbp.

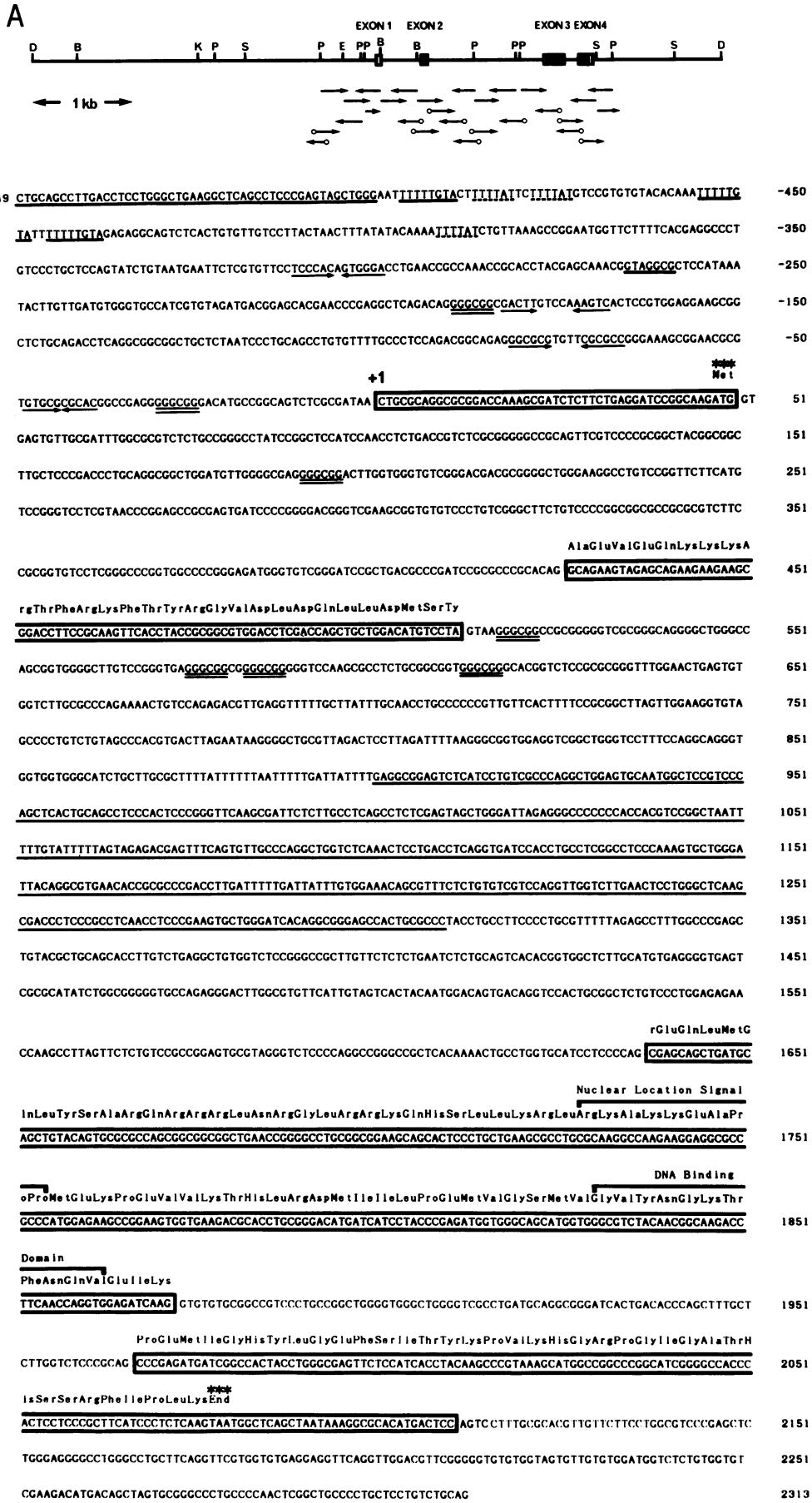

FIG. 2. *(Legend appears at the bottom of the opposite page.)*

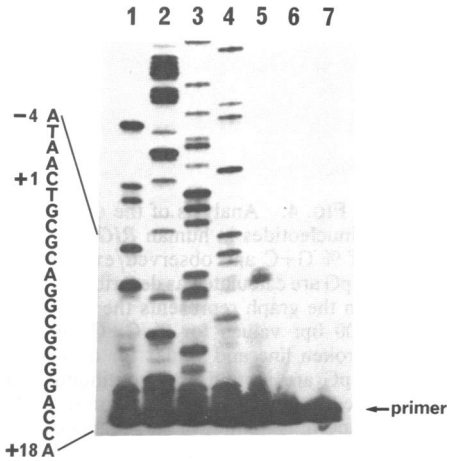

FIG. 3. Mapping of the transcription initiation site by primer extension analysis. Primer extension analysis was performed as described (19). A synthetic 18-mer (5′-CTCAGAAGAGATCGC-TTT-3′) complementary to the 5′ end of human *RIG* cDNA (4) was 5′ end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The labeled primer was hybridized to esophageal carcinoma RNA and extended by reverse transcriptase. Dideoxy sequencing reaction using the same 18-mer primer described above was performed to align the extended fragment. Primer-extended products and dideoxy sequencing reaction products were electrophoresed on a 5% polyacrylamide/8.3 M urea gel. Lanes 1, 2, 3, and 4, A, C, G, and T of a sequencing ladder, respectively; lane 5, primer extension experiment; lane 6, primer extension control experiment without RNA; lane 7, labeled primer. Three DNA bands differing in length by one nucleotide each were detected (lane 5) and the cytosine residue corresponding to the largest band was defined as the transcription initiation site and labeled +1. The nucleotide sequence of the sense strand and nucleotide numbers are indicated at the left.

very little suppression of this dinucleotide in the 5′ region of the gene. The further upstream portion of the 5′-flanking region (nucleotides −549 to −202) and the middle portion of intron 2 containing an *Alu* sequence (nucleotides 665–1415) had relatively low G+C contents (46% and 55%, respectively). The G+C content in the 3′ region of the gene containing exons 3 and 4 and intron 3 (nucleotides 1416–2313) was 63%.

Human *RIG* Pseudogenes. As shown in Fig. 1*C*, Southern blot analysis of *Dra* I-digested DNA showed several positive bands that hybridized to human *RIG* cDNA at stringent criterion, suggesting that human genomic *RIG* consists of a multigene family and/or that several copies of the pseudogenes exist in the human genome. We constructed a DNA library with *Sau*3AI partially digested esophageal carcinoma DNA and λEMBL3 and screened it with human *RIG* cDNA as probe. Twenty-one positively hybridized clones were plaque-purified and the phage DNAs were digested with *Dra* I. Southern blot hybridization revealed that the DNA library contained five types of independent clones and that these clones contained DNA fragments of 4.0, 3.2, 1.8, 1.0, and 0.8 kbp, as seen in Fig. 1*C*. Another genomic library, constructed with the 9- to 23-kbp *Dra* I fragments of the esophageal carcinoma DNA and λEMBL4, was found to contain an additional clone that corresponded to the 14-kbp *Dra* I fragment shown in Fig. 1*C*. Thus, all of the members of the *RIG* family that appeared on the Southern blots have now been accounted for.

To see the nature of the members of the *RIG* family, entire nucleotide sequences were determined for genomic clones 1.8 and 1.0, which contained 1.8-kbp and 1.0-kbp *Dra* I fragments, respectively, and partial sequences of genomic clones 4.0, 3.2, and 0.8, which contained 4.0-, 3.2-, and 0.8-kbp *Dra* I fragments, respectively, were also obtained (data not shown). Although clones 4.0, 3.2, 1.8, and 0.8 had nucleotide sequences homologous to human *RIG* cDNA, they did not contain any introns and possessed a stretch of adenylic acid residues shortly downstream from the poly(*A*) consensus sequence, all of which are characteristic of the processed type of pseudogene (27, 28). Clone 1.0 had a 180-bp sequence highly homologous (87%) to nucleotide sequence 133–312 of the *RIG* cDNA (4). Clone 14, which contained the 14-kbp *Dra* I fragment, was regarded as a pseudogene or a related sequence, because some oligonucleotide primers specific to human *RIG* cDNA did not work on the insert of this clone in the dideoxy sequencing reactions.

DISCUSSION

Several copies of *RIG* found in the human genome have been characterized. One has the characteristics of a functional gene, and six have been identified as pseudogenes. The functional gene has four exons and three introns, and the nucleotide sequences at the exon–intron boundaries are similar to the consensus sequences found in many eukaryotic gene sequences (24). The nucleotide sequences of the exons are identical to the *RIG* cDNA sequences (4). A single size class of mRNA (0.7 kb) was detected by RNA blotting (Fig. 1*A*). When the same filter in Fig. 1*B* was hybridized using the *Bam*HI fragment of intron 1 of human *RIG* (Fig. 2*A*) as probe, only a single band was detected in each digest (data not shown). These results indicate that there is a single copy of the functional *RIG* in a haploid set of the human genome and that *RIG* mRNA is transcribed from this gene.

In the 5′-flanking region of human *RIG* there is no typical TATA or CAAT sequence. This is consistent with the 5′ sequences of housekeeping genes, such as hypoxanthine phosphoribosyltransferase (25), adenosine deaminase (37), and dihydrofolate reductase genes (29, 30). The most characteristic feature of the gene is its extraordinarily high GC content. The 5′-flanking region, exon 1, intron 1, exon 2, and the 5′ portion of intron 2 are particularly rich in G+C, and these regions contain "CpG islands" (Fig. 4), regions of DNA with a high GC content and a high frequency of CpG dinucleotides relative to the bulk genome (26, 31, 32). Seven copies of GGGCGG, a basic motif of Sp1-binding sites (33), are found in the CpG islands at positions −189 to −184, −30 to −25, 191–196, 517–522, 575–580, 583–588, and 613–618 (Fig. 2*B*), and the second sequence, GGGGCGGGAC, at 30 bp upstream from the transcription initiation site is completely identical with the consensus sequence of Sp1-binding sites (34, 35), suggesting that expression of the gene may be regulated by binding of the transcription factor Sp1. Therefore, as the gene is expressed in all tissues and cells so far

FIG. 2 (*on opposite page*). (*A*) Restriction map and sequencing strategy of human *RIG*. The direction of transcription is from left to right. Restriction sites are shown on the top: B, *Bam*HI; D, *Dra* I; E, *Eco*RI; K, *Kpn* I; P, *Pst* I; S, *Sac* I. Exons are indicated by open boxes (untranslated regions) and by closed boxes (protein coding regions). Arrows indicate the direction and the extent of sequence determination. Open circles indicate synthetic primers. (*B*) Nucleotide sequence of human *RIG*. The nucleotide sequence of the sense strand is shown. Nucleotide residues are numbered in the 5′ → 3′ direction, beginning with the transcription initiation site, and nucleotides on the 5′ side of residue 1 are indicated by negative numbers. Exons are indicated by thick solid line boxes. The ATG codon for initiator methionine and the TAA termination codon are indicated by asterisks. The amino acid sequence is shown above the nucleotide sequence. GC boxes are indicated by double underlines. Direct repeats are indicated as follows: TTTTGTA, thin underline; TTTTAT, dotted underline. Inverted repeats are indicated by an underline with arrowhead. Sequences homologous to *Alu* consensus sequence (23) are indicated by underline.

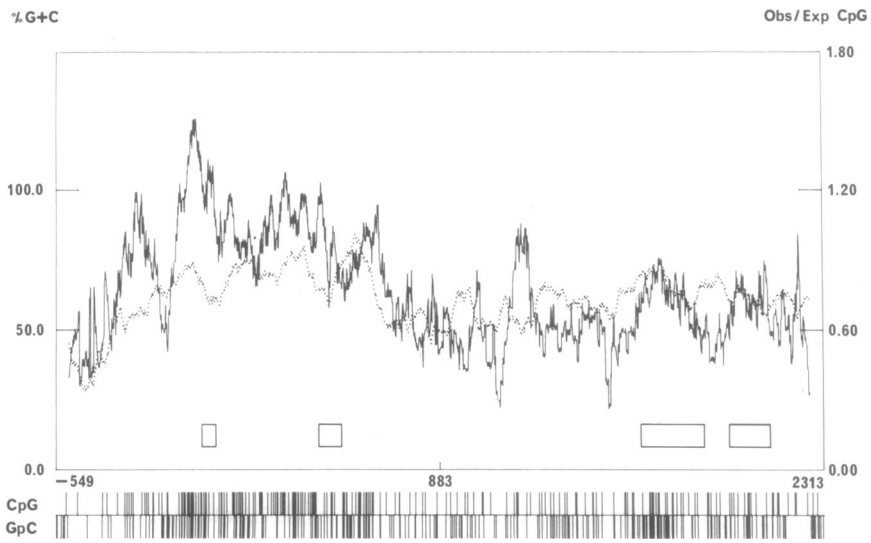

FIG. 4. Analysis of the distribution of CpG dinucleotides in human *RIG*. Moving averages of % G+C and observed/expected (Obs/Exp) CpG are calculated as described (26). Each point on the graph represents the average values for 100 bp: values for % G+C are plotted as a broken line and values for observed/expected CpG are plotted as a continuous line. Underneath the graph, the position of each CpG and GpC dinucleotide in the sequence is indicated by a vertical line: exons are marked by open boxes. Nucleotide numbers are shown on the abscissa. Analysis was carried out using the GENETYX program (Software Development, Tokyo).

examined (refs. 1 and 4–8; unpublished work), it is attractive to speculate that *RIG* belongs to the class of "housekeeping" genes whose products are necessary for the growth of all cell types. The presence of direct repeats in housekeeping promoters has been reported (36). As shown in Fig. 2*B*, several kinds of direct repeat and inverted repeat are also found in the 5′-flanking region of the gene. It should be noted here that there are two *Alu* sequences (nucleotides −549–−500 and 901–1310) and that the CpG islands are sandwiched between these *Alu* sequences.

The occurrence of processed pseudogenes appears to be a general phenomenon, especially in mammalian housekeeping genes (27, 28). The large number of processed *RIG* pseudogenes provides additional support for our speculation that *RIG* belongs to the class of housekeeping genes.

This work has been supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

1. Takasawa, S., Yamamoto, H., Terazono, K. & Okamoto, H. (1986) *Diabetes* **35,** 1178–1180.
2. Yamagami, T., Miwa, A., Takasawa, S., Yamamoto, H. & Okamoto, H. (1985) *Cancer Res.* **45,** 1845–1849.
3. Okamoto, H. (1985) *BioEssays* **2,** 15–21.
4. Inoue, C., Shiga, K., Takasawa, S., Kitagawa, M., Yamamoto, H. & Okamoto, H. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6659–6662.
5. Shiga, K., Inoue, C., Takasawa, S., Yonekura, H., Yamamoto, H. & Okamoto, H. (1986) *Seikagaku* **58,** 722 (abstr.).
6. Takasawa, S., Inoue, C., Shiga, K. & Kitagawa, M. (1990) in *Molecular Biology of the Islets of Langerhans,* ed. Okamoto, H. (Cambridge Univ. Press, Cambridge, U.K.), pp. 287–299.
7. Inoue, C., Igarashi, K., Kitagawa, M., Terazono, K., Takasawa, S., Obata, K., Iwata, K., Yamamoto, H. & Okamoto, H. (1988) *Biochem. Biophys. Res. Commun.* **150,** 1302–1308.
8. Okamoto, H., Yamamoto, H., Takasawa, S., Inoue, C., Terazono, K., Shiga, K. & Kitagawa, M. (1988) in *Frontiers in Diabetes Research: Lessons from Animal Diabetes II,* eds. Shafrir, E. & Renold, A. E. (Libbey, London), pp. 149–157.
9. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
10. Ohsawa, K., Hayakawa, Y., Nishizawa, M., Yamagami, T., Yamamoto, H., Yanaihara, N. & Okamoto, H. (1985) *Biochem. Biophys. Res. Commun.* **132,** 885–891.
11. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12,** 7035–7056.
12. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 280–281.
13. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
14. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3683–3687.
15. Yonekura, H., Nata, K., Watanabe, T., Kurashina, Y., Yamamoto, H. & Okamoto, H. (1988) *J. Biol. Chem.* **263,** 2990–2997.
16. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
17. Terazono, K., Yamamoto, H., Takasawa, S., Shiga, K., Yonemura, Y., Tochino, Y. & Okamoto, H. (1988) *J. Biol. Chem.* **263,** 2111–2114.
18. Kaiser, K. & Murray, N. E. (1985) in *DNA Cloning: A Practical Approach,* ed. Glover, D. M. (IRL, Oxford), Vol. 1, pp. 1–47.
19. Yamagami, T., Ohsawa, K., Nishizawa, M., Inoue, C., Gotoh, E., Yanaihara, N., Yamamoto, H. & Okamoto, H. (1988) *Ann. N.Y. Acad. Sci.* **527,** 87–102.
20. Helms, C., Graham, M. Y., Dutchik, J. E. & Olson, M. V. (1985) *DNA* **4,** 39–49.
21. Hattori, M. & Sakaki, Y. (1986) *Anal. Biochem.* **152,** 232–238.
22. Mizusawa, S., Nishimura, S. & Seela, F. (1986) *Nucleic Acids Res.* **14,** 1319–1324.
23. Jelinek, W. R. & Schmid, C. W. (1982) *Annu. Rev. Biochem.* **51,** 813–844.
24. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. & Sharp, P. A. (1986) *Annu. Rev. Biochem.* **55,** 1119–1150.
25. Patel, P. I., Framson, P. E., Caskey, C. T. & Chinault, A. C. (1986) *Mol. Cell. Biol.* **6,** 393–403.
26. Gardiner-Garden, M. & Frommer, M. (1987) *J. Mol. Biol.* **196,** 261–282.
27. Hollis, G. F., Hieter, P. A., McBride, O. W., Swan, D. & Leder, P. (1982) *Nature (London)* **296,** 321–325.
28. Lee, M. G.-S., Lewis, S. A., Wilde, C. D. & Cowan, N. J. (1983) *Cell* **33,** 477–487.
29. Chen, M.-J., Shimada, T., Moulton, A. D., Cline, A., Humphries, R. K., Maizel, J. & Nienhuis, A. W. (1984) *J. Biol. Chem.* **259,** 3933–3943.
30. Yang, J. K., Masters, J. N. & Attardi, G. (1984) *J. Mol. Biol.* **176,** 169–187.
31. Bird, A. P. (1986) *Nature (London)* **321,** 209–213.
32. Dynan, W. S. (1986) *Trends Genet.* **2,** 196–197.
33. Dynan, W. S. & Tjian, R. (1983) *Cell* **35,** 79–87.
34. Briggs, M. R., Kadonaga, J. T., Bell, S. P. & Tjian, R. (1986) *Science* **234,** 47–52.
35. Johnson, A. C., Ishii, S., Jinno, Y., Pastan, I. & Merlino, G. T. (1988) *J. Biol. Chem.* **263,** 5693–5699.
36. McGrogan, M., Simonsen, C. C., Smouse, D. T., Farnham, P. J. & Schimke, R. T. (1985) *J. Biol. Chem.* **260,** 2307–2314.
37. Valerio, D., Duyvesteyn, M. G. C., Dekker, B. M. M., Weeda, G., Berkvens, Th. M., van der Voorn, L., van Ormondt, H. & van der Eb, A. J. (1985) *EMBO J.* **4,** 437–443.