The human c-Ha-*ras*2 is a processed pseudogene inactivated by numerous base substitutions

Jun Miyoshi, Masaaki Kagimoto*, Ei-ichi Soeda+ and Yoshiyuki Sakaki

Research Laboratory for Genetic Information, *First Department of Internal Medicine, Kyushu University, Faculty of Medicine, Fukuoka 812, and +National Institute of Genetics, Mishima, Shizuoka 411, Japan

Received 14 December 1983; Accepted 19 January 1984

ABSTRACT

The human c-Ha-ras2 gene, one of two known members of the Harvey ras family, is reportedly located on the X-chromosome and has lost introns (1, 2). There has heretofore been no information on its precise gene structure and oncogenic potential. We have determined the nucleotide sequence of the c-Ha-ras2 and demonstrate that it is a processed pseudogene surrounded by several direct repeats and contains numerous base substitutions as well as a notable mutation (AGT at codon 12 of the p21 protein) responsible for oncogenic conversion of the known ras genes (3-8).

INTRODUCTION

Normal human DNA contains sequences related to retroviral oncogenes, and activation of these proto-oncogenes is considered to be involved in expression of the transformed phenotype in naturally occurring tumors. The transfection assay of NIH 3T3 cells with tumor DNA disclosed activated versions of ras-related proto-oncogenes (9-16), comprised of at least three members (c-Ha-ras, c-Ki-ras and N-ras) (17), in human solid tumors, cell lines of epithelial origin and certain types of hematopoietic malignancies (18, 19). Based on hybridization analysis, the human genome was found to contain additional cellular homologues of transforming genes of the Harvey and Kirsten murine sarcoma viruses (1). As characterization of normal ras-related sequences and determination of their exact numbers might be of clinical significance in the diagnosis and prevention of cancer, we attempted to identify homologous sequences by screening a human genomic DNA library using the v-Ha-ras probe at various hybridization stringency. Among these clones, we report here the nucleotide sequence of the c-Ha-ras2 and demonstrated that it is a processed and inactivated pseudogene accompanied by several direct repeats.

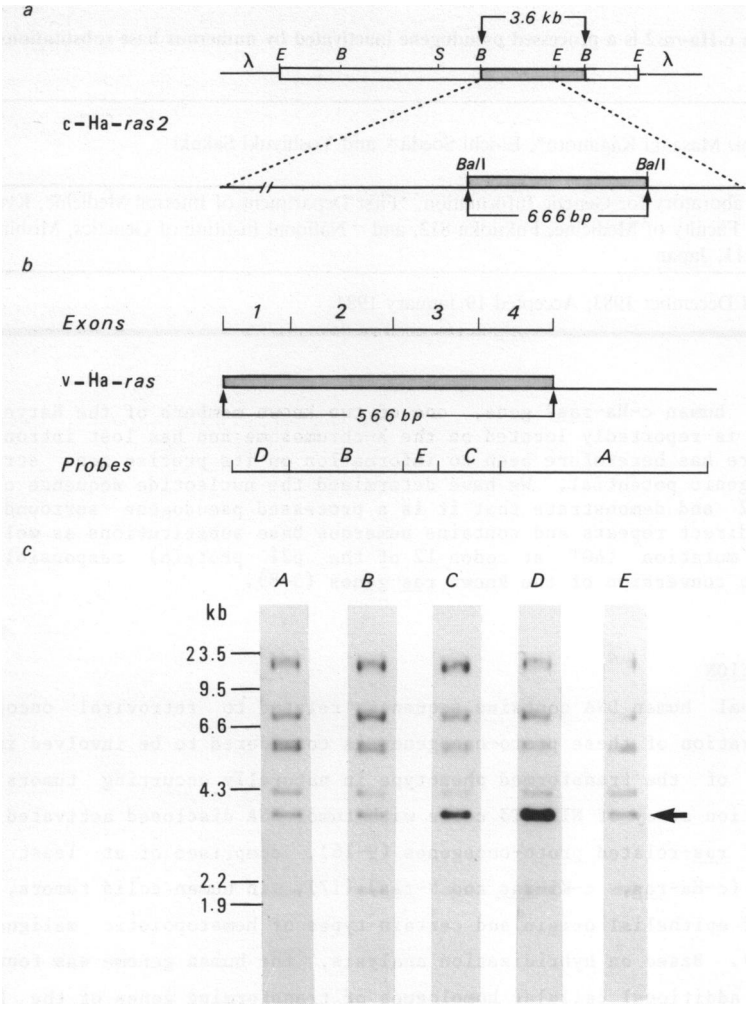

Fig. 1. Relationship between the c-Ha-ras2 and the oncogene of Harvey-MSV (v-Ha-ras).

a. Schematic representation of the c-Ha-ras2 gene. The human DNA insert is boxed. Shaded boxes indicate fragments homologous to the v-Ha-ras. Abbreviations of restriction enzymes are as follows; E; EcoRI, B; BamHI, S; SacI.

b. Digestion of the 670-bp HindIII-PstI fragment of the v-Ha-ras gene with HinfI produces five fragments (A-E), and the correlation to four coding exons (1-4) of the v-Ha-ras gene is shown schematically.

c. The recombinant λ DNAs containing the c-Ha-ras2 gene were digested with BamHI, electrophoresed in agarose (0.8% w/v) gels, blotted to nitrocellulose filters and hybridized in condition of 6X SSC and 55°C for 20 h. to nick-translated probes; five probes listed in b. were used in each lane from A to E. Filters were then washed four times in 0.1 x SSC

at 55°C and exposed to Fuji Medical X-ray film at -70°C with intensifying screens for 24 h. The arrow marks the 3.6-kb BamHI fragment specifically hybridize with each probe. Nonspecific hybridization backgrounds of λ arms and human DNA sequences were also visualized in this low stringent condition. DNA fragments of HindIII-digested λ served as size standards.

MATERIALS AND METHODS

Preparation of DNA probes

The ras probes were prepared by digestion of the Harvey MuSV sequences cloned at the EcoRI site of pBR322 with appropriate restriction enzymes, isolated on 1 % agarose or 3.5-5 % acrylamide gels, and recovered by electroelution. All probes were labeled with [α-^{32}P]dCTP by nick-translation to a specific activity of ~$2x10^8$ c.p.m. per μg DNA, and subjected to plaque hybridization and Southern blot analysis (20, 21). Enzymes were purchased from Takara Shuzo Co., Ltd., Japan, and radio-isotopes were from New England Nuclear.

Isolation of human v-ras homologous sequences

The human genomic DNA library, obtained from T. Maniatis, was screened for human v-ras homologous sequences by the method of Benton and Davis (20). The c-Ha-ras2 clone was isolated at the standard stringent condition as described (22).

DNA sequencing

DNA sequencing was performed entirely by the partial chemical cleavage method of Maxam and Gilbert (23). The resulting fragments were resolved on 0.3 mm thin 8 % and 20 % acrylamide/8 M urea gels. All sequences presented were obtained from both DNA strands and/or in both directions.

RESULTS AND DISCUSSION

We isolated a recombinant λ Charon 4A phage containing a 14.2-kb EcoRI insert of human DNA which exhibited a 3.6-kb internal BamHI fragment that specifically hybridized with the v-Ha-ras probe. This 3.6-kb BamHI fragment was subsequently used for restriction enzyme analysis and proved to be the clone identical to the c-Ha-ras2 reported by Chang et al (1) (Fig. 1a). The structural relationship between the c-Ha-ras2 and the v-Ha-ras sequences was studied by Southern blot analysis using ras probes divided by HinfI in the low stringent condition as described in the Fig. 1 legend. As shown in Fig. 1b and 1c, the 3.6-kb BamHI fragment produced strong hybridization with the 99-bp HinfI fragment D, and also hybridized with the 110-bp HinfI fragment C, but not with HinfI fragments A, B and E. These findings suggest that the

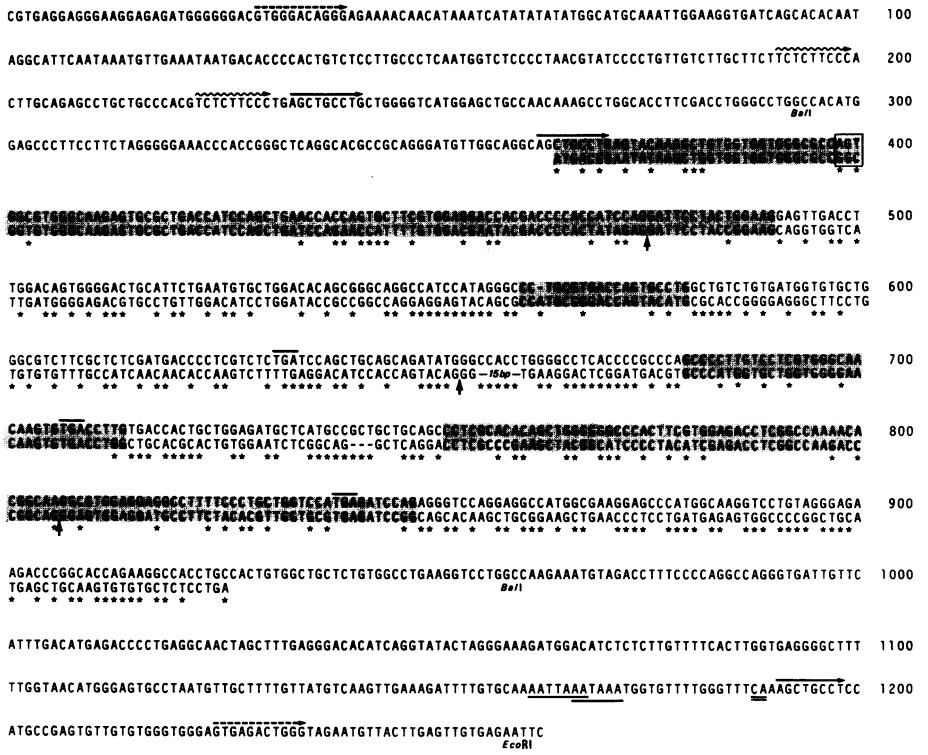

Fig. 2. Comparative nucleotide sequence analysis of the human c-Ha-ras2 and c-Ha-ras1 proto-oncogenes.

The lower line shows the sequence of exons of the c-Ha-ras1 reported by Capon et al (14). Homologies (~80%) between the two are indicated by the shaded boxes. Dashes represent insertions introduced for alignment to maximize homology, and asterisks mark differences in the two sequences. Arrows show the positions of RNA splicing in the c-Ha-ras1 gene. Direct repeats are overlined by solid, wavy and broken-lined arrows, in-frame termination codons are overlined, and underlined sequences are the poly(A) addition signal and the poly(A) addition site. These features are discussed in the text.

c-Ha-ras2 shares extensive homology with the first exon of the ras gene, and rather weak homology with the junction of the third and fourth exons, and very little homology with the second and 5'-half of the third exon. Further analysis revealed that the homologous region was confined to the 666-bp BalI fragment (Fig. 1a).

The DNA sequence covering the entire c-Ha-ras2 was determined by the Maxam-Gilbert method and is presented in Fig. 2. Comparative sequence an-

alysis of the c-Ha-ras2 and the v-Ha-ras/ c-Ha-ras1 (3, 24) demonstrated that the homologous sequence is defined between positions 365 and 926. Distribution of homologous segments fundamentally parallels the findings of Southern blot analysis. All the nucleotides and deduced amino acids, between the third codon for glutamate and the 23rd codon for leucine in the first exon of p21 (position 371-433), coincide with the v-Ha-ras/c-Ha-ras1 sequence with the several exceptions. Although this homologous segment ($\backsim$80% homology) continues downstream to the first splicing point of the c-Ha-ras1 and extends further 15-bp 3' to the first exon (position 365-490), numerous amino acid changes were observed below the 24th codon for asparagine. Further comparison reveals three short stretches of $>$ 75% homology at positions 561-579, 680-714, and 752-848. The last one also spans the third splicing point. These two homologous segments spanning junctions of the first-second and third-fourth exons seem to be consistent with a property of the c-Ha-ras2 as a spliced or processed gene, despite its affection by numerous base substitutions.

We found no obvious sequences of transcription initiation signals, RNA cap sites and ATG translation initiation codons indispensable to ordinary eukaryotic genes. Several ATG codons are present in the 5' region to the protein coding-like sequences, but these are of no significance considering that three TGA termination codons are present at positions 632, 707 and 839 (overlined in Fig. 2) in the same reading frame aligned with the p21 initiation codon of the v-Ha-ras and the c-Ha-ras1. These findings strongly suggest that the c-Ha-ras2 is biologically silent and that it would not be rescued by insertion of promoter or enhancer like elements. Furthermore, we found evidence to support the idea that the c-Ha-ras2 is probably translocated via an RNA intermediate. Doublet of poly(A) addition signals AATTAAA and AATAAA (25, 26) is located between positions 1162 and 1172 (underlined in Fig. 2), 236-bp downstream from the end of the coding-like region, and the site of poly(A) addition is present 15-bp from the poly(A) addition signals (double-underlining in Fig. 2). It is notable that the sequence AGCTGCCT flanking the poly(A) addition site (position 1191) is also present just at the beginning of the c-Ha-ras2 (position 363) followed by the third codon GAG, thus it forms 8-bp direct repeats just outside the c-Ha-ras2. The sequence AGCTGCCT(G) is also found at position 234, but this one does not seem to be directly involved in generation of the c-Ha-ras2 because the region between positions 234 and 363 is not homologous to the 5' untranslated region of the c-Ha-ras1. Although the c-Ha-ras2 deletes most of the poly(A)

tail, it is quite conceivable that this gene is a processed pseudogene accompanied by direct repeats probably generated during the integration step. The existence of additional direct repeats GTG$^{G}_{A}$GAC$^{A}_{T}$GGG (at positions 30 and 1225) and TCTCTTCCC (at positions 191 and 223) surrounding the c-Ha-ras2 might reflect recombination events which have occurred in the process of DNA rearrangement.

Genetic mechanisms including gene amplification (c-myc and c-Ki-ras) (12, 27, 28), insertion of retroviral controlling sequences (c-myc) (29-31) and a transposable element (c-mos) (32), translocation (c-myc and c-abl) (33-36) and spontaneous mutation (c-Ha-ras1 and c-Ki-ras2) (5-8) have been proposed to explain activation of proto-oncogenes. However, none of them is expected to convert the c-Ha-ras2 itself to an oncogenic form, considering its numerous base substitutions. The actual role of the c-Ha-ras2 in human carcinogenesis is thus unknown.

It might be of some interest that the c-Ha-ras2 has the same mutation AGT at codon 12 of the p21 protein as the v-Ki-ras (4). The current concept is that a glycine residue at position 12 is important in normal p21 function (37) and that any mutation affecting the coding properties of codon 12 of the c-ras proto-oncogene will lead to its malignant activation (38). Although oncogenic potential of the c-Ha-ras2 is not convincing, it might be possible that the extensively homologous sequences (positions 371-433) containing the critical mutation of serine at codon 12 of p21 are involved in somatic recombination like gene conversion (39) with other members of the c-Ha-ras gene family and result in the formation of an oncogenic hybrid gene.

ACKNOWLEDGEMENTS

We thank K. Shimizu for many useful suggestions, M. Ohara for comments on the manuscript, and Y. Takagi for generous support and encouragement. This work was supported by grants from the Ministry of Education, Culture, and Science of Japan and the Fukuoka Cancer Society.

REFERENCES

1. Chang, E. H., Gonda, M. A., Ellis, R. W., Scolnick, E. M. and Lowy, D. R. (1982) Proc. Natl. Acad. Sci. USA 79, 4848-4852.
2. O'Brien, S. J., Nash, W. G., Goodwin, J. L., Lowy, D. R. and Chang, E. H. (1983) Nature 302, 839-842.
3. Dhar, R., Ellis, R. W., Shih, T. Y., Oroszlan, S., Shapiro, B., Maizel, J., Lowy, D. and Scolnick, E. (1982) Science 217, 934-937.
4. Tsuchida, N., Ryder, T. and Ohtsubo, E. (1982) Science 217, 937-939.
5. Tabin, C. J., Bradley, S. M., Bargmann, C. I., Weinberg, R. A.,

Papageorge, A. G., Scolnick, E. M., Dhar, R., Lowy, D. R. and Chang, E. H. (1982) Nature 300, 143-149.
6. Reddy, E. P., Reynolds, R. K., Santos, E. and Barbacid,, M. (1982) Nature 300, 149-152.
7. Taparowsky, E., Suard, Y., Fasano, O.,, Shimizu, K., Goldfarb, M., and Wigler, M. (1982) Nature 300, 762-765.
8. Shimizu, K., Birnbaum, D., Ruley, M. A., Fasano, O., Suard, Y., Edlund, L., Taparowsky, E., Goldfarb, M. and Wigler, M. (1983) Nature 304, 497-500.
9. Parada, L. F., Tabin, C. J., Shih, C. and Weinberg, R. A. (1982) Nature 297, 474-478.
10. Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S. and Barbacid, M. (1982) Nature 298, 343-347.
11. Der, C. J., Krontiris, T. G., and Cooper, G. M. (1982) Proc. Natl. Acad. Sci. USA 79, 3637-3640.
12. McCoy, M. S., Toole, J. J., Cunningham, J. M., Chang, E. H., Lowy, D. R. and Weinberg, R. A. (1983) Nature 302, 79-81.
13. Shimizu, K., Birnbaum, D., Ruley, M. A., Fasano, O., Suard, Y., Edlund, L., Taparowsky, E., Goldfarb, M. and Wigler, M. (1983) Nature 304, 497-500.
14. Capon, D. J., Seeburg, P. H., McGrath, J. P., Hayflick, J. S., Edman, U., Levinson, A. D. and Goeddel, D. V. (1983) Nature 304, 507-513.
15. Shimizu, K., Goldfarb, M., Perucho, M. and Wigler, M. (1983) Proc. Natl. Acad. Sci. USA 80, 383-387.
16. Hall, A., Marshall, C. J., Spurr, N. K. and Weiss, R. A. (1983) Nature 303, 396-400.
17. Shimizu, K., Goldfarb, M., Suard, Y., Perucho, M., Li, Y., Kamata, T., Feramisco, J., Stavnezer, E., Fogh, J. and Wigler, M. H. (1983) Proc. Natl. Acad. Sci. USA 80, 2112-2116.
18. Pulciani, S., Santos, E., Lauver, A. V., Long, L. K., Aaronson, S. A. and Barbacid, M. (1982) Nature 300, 539-542.
19. Westin, E. H. et al. (1982) Proc. Natl. Acad. Sci. USA 79, 2490-2494.
20. Benton, W. D. and Davis, R. W. (1977) Science 196, 180-182.
21. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
22. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. and Efstratiadis, A. (1978) Cell 15, 687-701.
23. Maxam, A. and Gilbert, W. (1980) Meth. Enzym. 65, 499-580.
24. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. and Goeddel, D. V. (1983) Nature 302, 33-37.
25. Proudfoot, N. J. and Brownlee, G. G. (1976) Nature 263, 211-214.
26. Unterman, R. D., Lynch, K. R. Nakhasi, H. L., et al. (1981) Proc. Natl. Acad. Sci. USA 78, 3478-3482.
27. Collins, S. and Groudine, M. (1982) Nature 298, 679-681.
28. Dalla-Favera, R., Wong-Staal, F. and Gallo, R. C. (1982) Nature 299, 61-63.
29. Hayward, W. S., Neel, B. G. and Astrin, S. M. (1981) Nature 290, 475-480.
30. Payne, G. S., Bishop, J. M. and Varmus, H. E. (1982) Nature 295, 209-214.
31. Chang, E. H., Furth, M. E., Scolnick, E. M. and Lowy, D. R. (1982) Nature 297, 479-484.
32. Rechavi, G., Givol, D. and Canaani, E. (1982) Nature 300, 607- 610.
33. Dalla-Favera, R., Martinotti, S., Gallo, R. C., Ericson, J. and Croce, C. M. (1983) Science 219, 963-967.
34. Adams, J. M., Gerondakis, S., Weeb, E., Corcoran, C. M. and Cory, S. (1983) Proc. Natl. Acad. Sci. USA 80, 1982-1986.

35. Stanton, L. W., Watt, R. and Marcu, K. B. (1983) Nature 303, 401-406.
36. de Klein, A. et al. (1982) Nature 300, 765-767.
37. Wierenga, R. K. and Hol, W. G. J. (1983) Nature 302, 842-844.
38. Santos, E., Reddy, E. P., Pulciani, S., Feldmann, R. J. and Barbacid, M. (1983) Proc. Natl. Acad. Sci. USA 80, 4679-4683.
39. Lewin, R. (1983) Science 219, 1052-1054.