Vol. 166, No. 1, 1990 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1990 Pages 1-10

Transcription of Human Endogenous Retroviral Long Terminal Repeat (LTR) Sequence in a Lung Cancer Cell Line

Naohiro Tomita[1,2], Akira Horii[1,2], Sadayuki Doi[1,2], Hideoki Yokouchi[1,2], Michio Ogawa[2], Takesada Mori[2] and Kenichi Matsubara[1]*

[1] Institute for Molecular and Cellular Biology, Osaka University, Yamada-oka, Suita, 565 Japan

[2] Second Department of Surgery, Osaka University Medical School, Fukushima-ku, Osaka, 553 Japan

Received October 23, 1989

The human genome carries several endogenous retroviral sequences. One of them that we named 'HERV-A', carries almost the complete sequence of the long terminal repeat (LTR), and is located in the 5' region of the amylase genes (M.Emi, A.Horii, N.Tomita, T.Nishide, M.Ogawa, T.Mori and K.Matsubara, Gene 62: 229-235, 1988). Using this sequence as a probe, we found a 1.4 kb LTR transcript(s) in a lung cancer cell line. No corresponding transcript was observed in control cells. Two partial, but different cDNA clones were obtained, and each one was found to be a transcript starting within human sequences at 5' upstream from the LTR and ending within the LTR sequence.

Mammalian genomes carry multiple copies of endogenous retroviral sequences (1, 2, 3, 4), and those that are found in the human genome are usually called HERV sequences. Most of such HERV sequences have been detected by using probes prepared from exogenous retroviral sequences, such as murine leukemia virus or Moloney murine leukemia virus, but some of them have been detected while analyzing human genes. For example, Mager et al. detected a retrovirus-like element near to the human β-globin gene (5), Maeda found a similar element in the first intron of the human haptoglobin-related gene (6), and Emi et al. discovered such a sequence in the 5' flanking regions of the human salivary and pancreatic amylase genes (7). This sequence, which we tentatively

0006-291X/90 $1.50

call 'HERV-A', carries a complete LTR sequence that is 70~71% homologous with the 5' LTR of the HERV reported by Steel et al. (8). It is possible that in the human genome, there are several hundred copies of this sequence (7, 9).

In almost all the cases analysed, the endogenous retroviruses are not transcriptionally active in human tissues, and are considered to be nonpathogenic (10). However, their expression has been detected in few cases, such as in human placenta and a colon cancer cell line (11, 12), in human fetal liver (13), in chorionic villi (14, 15), in several cultured tumor cells (16), and in colon tumors (17).

Recently, Kato et al. showed by S1 mapping analysis that one of such transcripts may contain a read-through human cellular sequence hooked to the 3' side of the LTR (18), suggesting the possibility that some of the HERV sequences might activate the neighbouring human gene. An attractive possibility is that some human gene which is related to carcinogenesis might be activated by such a mechanism.

We started to examine if the HERV-A LTR might be expressed in some of the human tissues, and if so, whether it might lead to activation of some of the human genes. Our results show that in a particular lung carcinoma cell line, transcriptions of at least two human cellular sequences continue into the HERV-A LTR.

MATERIALS AND METHODS

Tissue samples and cell line

Human normal tissues (placenta, lung and colon) and malignant tissues (lung cancer and colon cancer) were obtained in surgical operations. They were immediately frozen in liquid nitrogen and stored at −70°C until use. A cell line originating from human lung adenocarcinoma was established by Izumi et al., (19) and was propagated in BALB/c nu/nu athymic mice.

Preparation of mRNA

Total cellular RNA was isolated as described (20), and mRNA was purified by repeated passages through an oligo (dT) cellulose column (Type 7, Pharmacia Co., Sweden).

Preparation of LTR probe and regional probes of LTR cDNA

A 369-bp *Hin*dIII-*Bst*NI fragment of cos P2 (21) was subcloned to pUC18 (22) and used as a probe for Southern and Northern blottings or cDNA library screening. This DNA fragment covers the entire LTR which lies in 5' flanking region of the human pancreatic amylase gene *(amy2)* (21). *Alu*I restricted subfragments of an LTR cDNA clone, λLTR8 (Fig. 2), were used as regional probes for Southern blottings in Fig. 6. ^{32}P-labeled probes (specific activity 4-8×10^8 cpm / μg) were prepared by the random priming method (23) using [α-^{32}P] dCTP (Amersham, 3000 Ci / m mol; 1Ci=37GBq).

Southern and Northern blot hybridizations

Ten μg of high molecular weight DNA was digested twice with an appropriate restriction enzyme (Takara Shuzo, Kyoto, Japan), electrophoresed in a 0.7% agarose gel and transferred to a nylon filter (Gene Screen Plus, NEN. USA) (24).

mRNA was denatured by heating at 65°C for 15 min in 50% (v/v) formamide, electrophoresed in a 1% agarose / 2.2M formaldehyde gel as described (25), and then transferred to a nylon filter.

After baking at 80°C for 1-2hr, the filters were pre-hybridized at 65°C for several hours in a sealed plastic bag containing 6×SSC (1×SSC is 0.15M NaCl, 15mM Na citrate), 1% SDS. Hybridization was carried out at 65°C overnight in 6×SSC, 1% SDS, 20μg / ml of heat denatured herring sperm DNA, containing ^{32}P-labeled probe (1-2×10^6 cpm / ml). After the hybridization, the filter was rinsed at room temperature in 2×SSC, 1% SDS, washed twice at 65°C in the above solution for 30 min and then rinsed at room temperature in 0.1×SSC. Autoradiography was performed for several hours or several days at −70°C with or without an intensifying screen.

Construction of cDNA library and screening for LTR clones

A cDNA library was constructed from mRNA prepared from lung cancer cells using phage vector λgt10 with *Eco*RI linkers (26). To screen for clones carrying the LTR sequence, plaque hybridization was performed using a nylon filter (Gene Screen Plus, NEN. USA) (27). The candidate phage clones were purified and the inserts were subcloned in pUC 18 for further analyses.

cDNA sequencing analysis

The pUC-subcloned fragments or their selected subfragments were recloned in M13 mp18 or mp19 (22) and then subjected to nucleotide sequencing analyses by the dideoxy chain termination method (28, 29). The computer program of Wilbur and Lipman (30) was used for homology search.

RESULTS AND DISCUSSION

To detarmine whether the HERV-A sequence is expressed, we examined mRNAs from human placenta and lung by Northern blot hybridization. No detectable transcripts were observed. In contrast to those, lung cancers and colon cancers showed smear hybridizing signal (data not shown), indicating that some of the HERV-A LTR sequences are transcribed in these neoplastic tissues. One lung adenocarcinoma cell line (19) which has been propagated in athymic mice displayed a clear 1.4 kb single band (Fig. 1).

To further examine the properties of the 1.4 kb transcript detected in the lung cancer cell line, a cDNA library was constructed using λgt 10 as a vector. Out of 1×10^6 clones, 62 clones showed positive signals with the HERV-A LTR probe, among which, 12 clones having long inserts were selected for further analyses. Restriction mapping and sequencing studies have revealed that the positive sequences consist of two distinct molecular species, both of which contain an almost complete LTR sequence. The restriction maps with λLTR8 and λLTR22, the two representative clones, are shown in Fig. 2. Among the rest of the

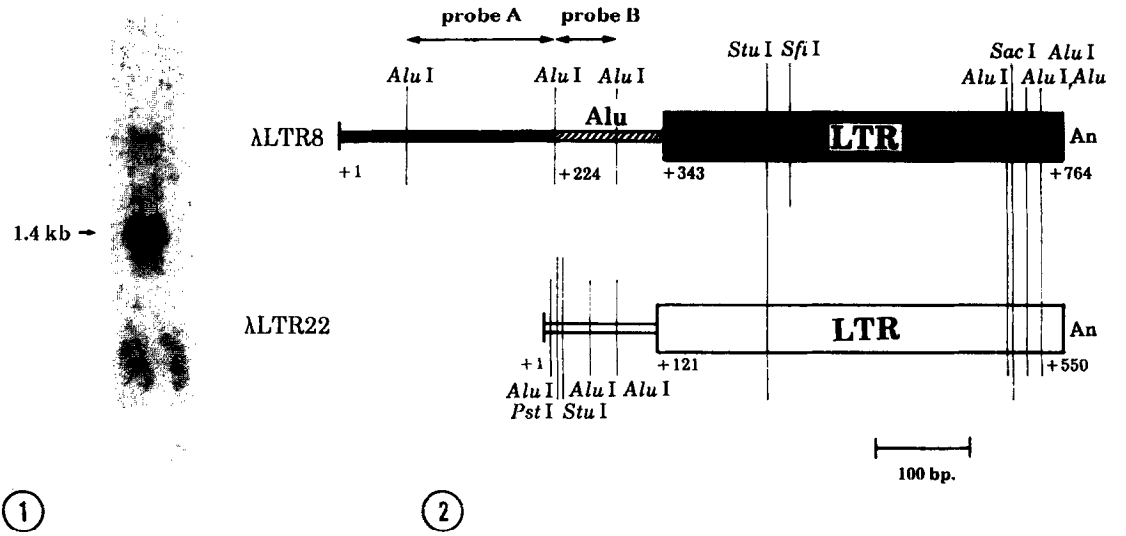

Fig. 1 Northern blot profile of mRNA from a lung cancer cell line using ^{32}P-labeled HERV-A LTR probe. Five μg of mRNA prepared from the cells propagated in nudemice was denatured, electrophoresed and subjected to Northern blotting using a 369-bp *Hin*dIII-*Bst*NI fragment of cos P2 as a probe. This DNA covers almost the entire region of the LTR sequence in HERV-A. For details, see 'MATERIALS AND METHODS'. Autoradiography was performed for 4 days at −70°C with an intensifying screen.

Fig. 2 Restriction maps of the two cDNA clones carrying the LTR transcripts isolated from cDNA library prepared from lung cancer cell line mRNA. Cleavage sites by some restriction enzymes are shown. The box represents the long terminal repeat (LTR) sequence. The numbers start from the 5' end of each type. An represents the poly A tail. The *Alu* homologous region of the clone λLTR8 is indicated. The regional probes of λLTR8 used in Southern blottings (see 'Fig. 6') are also indicated.

clones, five were the same as λLTR8, and the other five clones were the same as λLTR22. The sizes of the inserts of λLTR8 and λLTR22 are 764 bp and 550 bp, respectively, suggesting that both clones are likely to be partial clones. The slightly broad Northern blotting profile (Fig. 1) may reflect the presence of two major LTR transcripts. We attempted to prepare larger cDNA inserts without success. The reason for this failure is not clear.

The nucleotide sequences of the λLTR8 and λLTR22 inserts were examined, and the results are shown in Fig. 3. Homology search analysis showed that λLTR8 consists of three portions: A 223 bp (nt. +1~+223) region from the 5' end, having no homology with known sequences, inclusive of the retroviral sequences, and a following 119 bp region (nt. +224~+342) which shares 76% homology with human repetitive '*Alu*' consensus sequence (31), and a 442 bp region (nt. +343~+764) at the 3' end, having 85% homology with the HERV-A

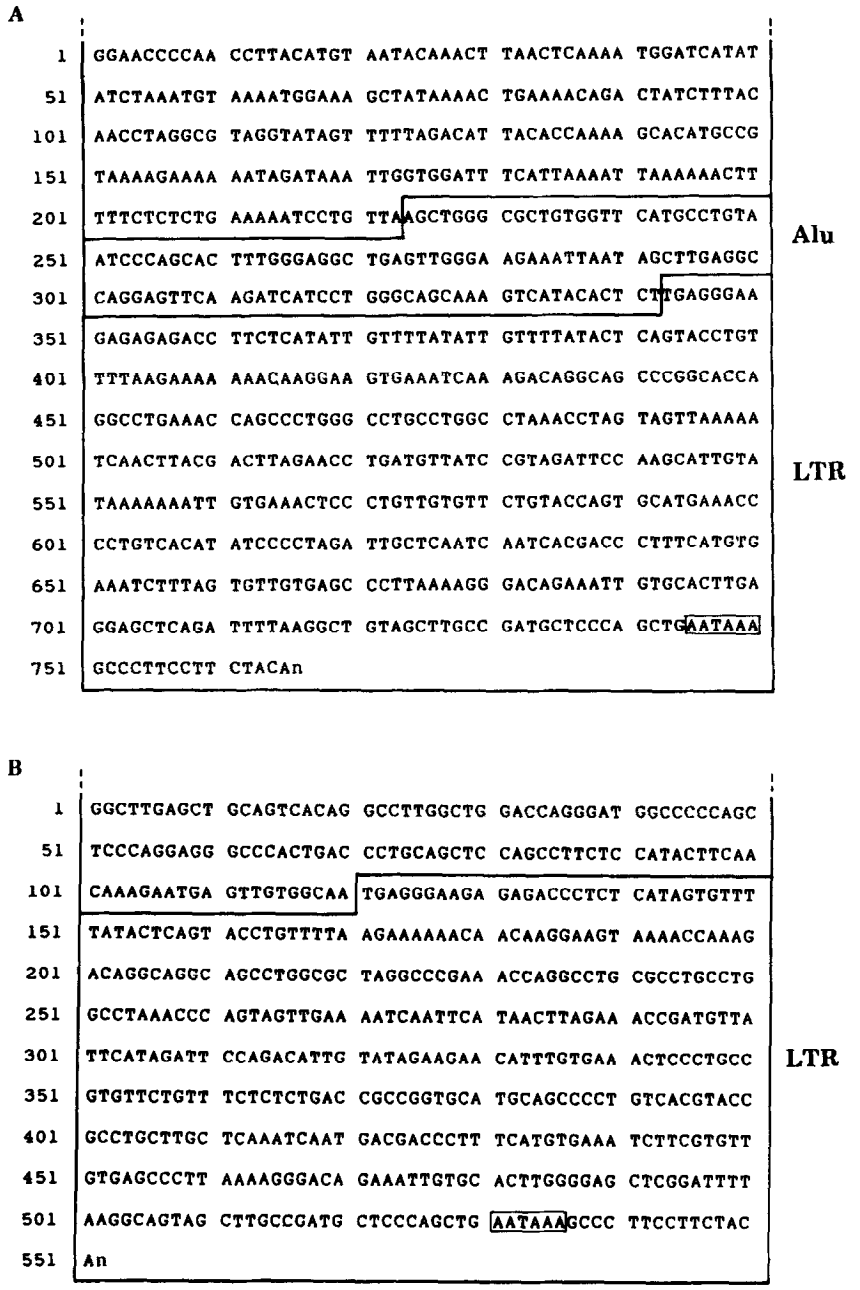

Fig. 3 Nucleotide sequences of the two cDNA clones carrying the LTR transcripts. A; λLTR8 and B; λLTR22. The sequence for λLTR8 is divided into three portions, i.e, 5' human cellular sequence region, *Alu* homologous region and the LTR region. The sequence for λLTR22 is divided into two portions, i.e, the possible 5' human cellular sequence region and the LTR region. The poly A stretch is represented by An, and the predicted poly A addition signal is boxed.

LTR sequence. The alignment of the *Alu*-like sequence to *Alu* consensus sequence is shown in Fig. 5. Southern blot analysis of *Eco*RI or *Pst*I digested human genomic DNA shows that the LTR probe prepared from cosP2 (21), gives

multiple bands (Fig. 6). By contrast, in the same test using the regional probe A prepared from the λLTR8 (see Fig. 2), only a few unique bands appeared, indicating that it represents the unique human cellular DNA sequence, unrelated to the retroviral sequence. In addition, the 67-bp *Alu* sequence (probe B) gave a smear hybridizing pattern, as expected. These observations show that the λLTR8 represents a transcript originating at some site in the cellular unique sequence, without splicing the region carrying the *Alu* repetitious sequence, and further reading into the LTR sequence. The transcription apparently terminated within the LTR sequence.

In another clone, λLTR22, a 430 bp region from the 3' end (nt. +121~+550) is 85% homologous with the LTR sequence in the HERV-A. This LTR sequence is hooked to a 120 bp region (nt. +1~+120) that has no nucleotide sequence homology with known retroviral or cellular sequences. The predicted amino acid sequence of this region, however, shows 39% homology with goat κ casein or 29% homology with human estrogen receptor (data not shown). Thus, the λLTR22 also represents a transcript originating within a human cellular sequence, reading into the LTR, and terminating there. The cellular sequence in the λLTR22 is only 120 bp in size, and was not appropriate for use as a probe for Southern blottings tests.

An 85% homology was observed between the LTR of λLTR8 and λLTR22, as shown in Fig. 4. Both LTR segments start with TG, as is seen with most of the integrated retroviral LTRs (32), and end with a poly A tail at 14-bp downstream from the putative poly A addition signal, AATAAA within the LTR. A *Sfi*I recognition sequence, GGCCNNNNNGGCC or its related sequence lies in both of the HERV-A LTR.

The implication of the presence of *Alu* sequences in the λLTR8 is not clear at present. It could be an exceptional transcript as been found in some mRNA's (33, 34). Alternatively, it could mean that the transcript has nothing to do with biological activities. Other possible activities can be imagined also with the RNA. Structurally, the consecutive presence of *Alu* and LTR in λLTR8 suggests strongly that a solo LTR or a 5' LTR of the HERV-A family may have selected the

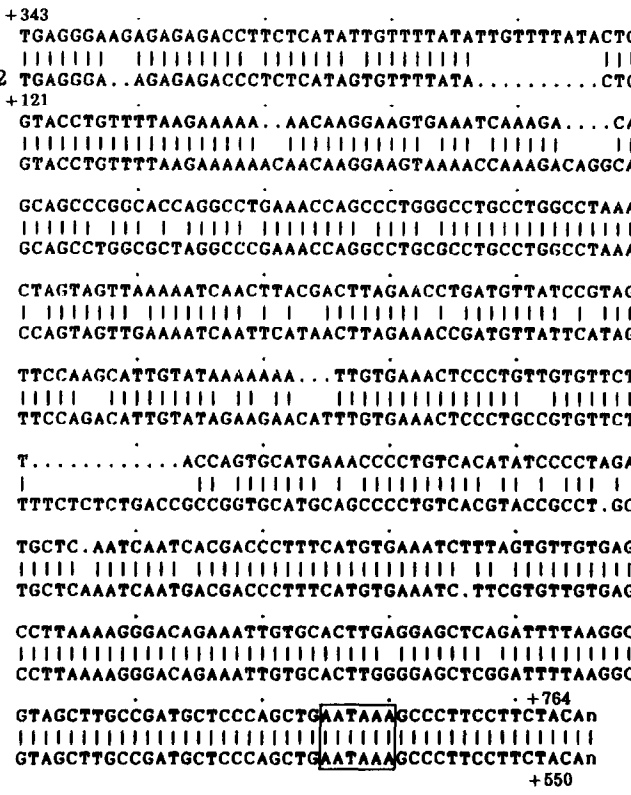

Fig. 4 Comparison of the LTR region of the two clones carrying LTR transcripts, λLTR8 and λLTR22. The LTR regions, (nt. +343~+764) of λLTR8 and (nt. +121~+550) of λLTR22 are aligned. The matching nucleotides are connected by vertical bars, and deletions are indicated by dots. The poly A stretch is represented by An, and the predicted poly A addition signal is boxed.

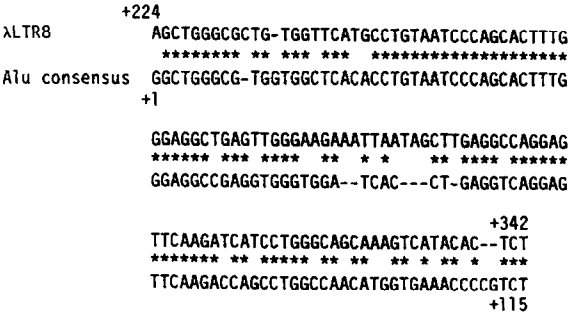

Fig. 5 Comparison of the sequences of the middle portion of the LTR8 and *Alu* consensus sequence. The sequence for a part of the λLTR8 (nt. +224~+342) and the *Alu* consensus sequence (nt. +1~+115, Deininger et al., 1981) are aligned. The matching nucleotides are indicated by asterisks.

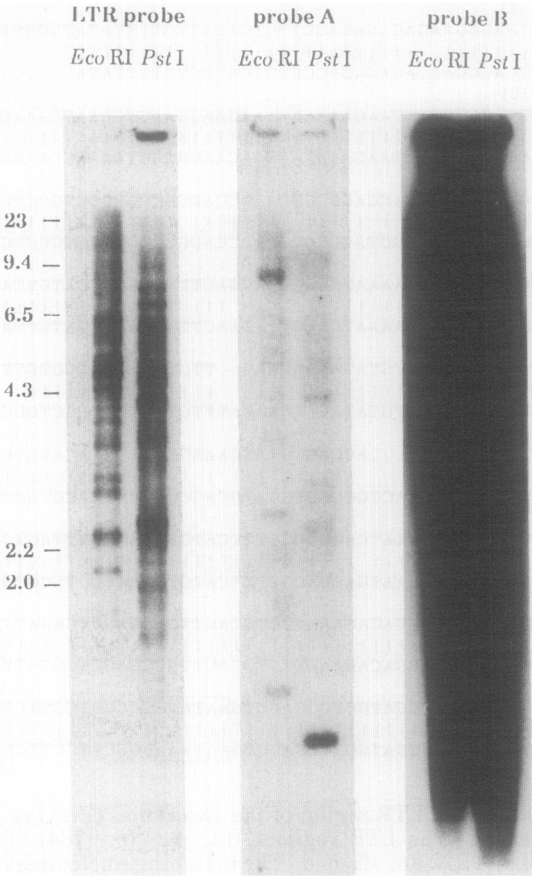

Fig. 6 Southern blot hybridizations of restricted human genomic DNA using the HERV-A LTR and two regional DNA's of λLTR8 as probes. Ten μg of human high molecular weight DNA were digested twice with the restriction enzymes indicated above each lane, electrophoresed and subjected to Southern blottings using the probes indicated on top of the figure. For details, see 'MATERIALS AND METHODS'. The probes used are as follows : LTR probe; A 369 bp-*Hin*dIII-*Bst*NI fragment of cos P2, probe A; A 154-bp *Alu*I-cleaved fragment of λLTR8, probe B; A 67-bp *Alu*I-cleaved fragment of λLTR8 (see 'Fig. 2'). Autoradiography was performed for 2 days for LTR probe and probe B, and for 7 days with an intensifying screen for probe A. Size markers (in kb) are the *Hin*dIII digest of λ phage DNA.

Alu sequence as the target for integration and landed in the middle of *Alu* sequences.

In the present study, we described the discovery of transcripts containing the HERV-A LTR sequence in a lung cancer cell line, although such transcripts are not detectable in normal lung tissue. To our knowledge, this is the first report on cloning and sequencing the cDNA's that carry the HERV-related sequence hooked to a human cellular sequence. Whether or not the expression of the LTR

sequence is related to carcinogenesis itself or to a process of establishing the cell line is a question to be answered in the future. Ono et al., demonstrated that in a human breast cancer cell line, expression of a HERV sequence (HERV-K) is stimulated by hormone (16), suggesting that some environmental changes surrounding the cell can affect the expression of these sequences. An interesting question in conjunction with expression of the HERV LTR is whether the transcription is initiated within the proviral genome or outside of it. The two distinct cDNA clones that we obtained from a lung cancer cell line were both of the latter type, viz. transcription starting somewhere in the human genome and reading into the LTR, although both of the clones were unfortunately incomplete. However, it shows that at least two genomic loci were used for initiation of new or augmented transcriptions in this cancer cell line. Recently, Kato et al. (18) have described three transcripts carrying HERV sequence, two of which are the transcripts initiated within the LTR sequence, and read through into human genomic sequences, although nucleotide sequences of these transcripts have not yet been shown. Our two cDNA clones are not of this type. It looks as though the transcripts described by Kato et al. have been made through activation of the promoter in LTR, whereas our transcripts have been made through activation of cellular promoters. An intriguing possibility is the enhancer-mediated activation of the cellular promoters in adjacent genes, and its transcription extending into the LTR, terminated using its poly A addition signal. Apart from this hypothesis , whether or not this LTR-mediated transcription has some function is the most intriguing question to promote further studies.

ACKNOWLEDGMENTS

We thank Miss K. Mimura for her excellent secretarial assistance. This work was supported by a Scientific Research Grant to K. M. from the Ministry of Education, Science and Culture, Japan.

REFERENCES

1. Todaro, G. J., Benveniste, R. E., Callahan, R., Lieber, M. M. and Sherr, C. J. (1974) Cold Spring Harbor Symp. Quant. Biol. 39, 1159-1168
2. Steffern, D. L. and Weinberg, R. (1978) Cell 15, 1003-1010

3. Martin, M. A., Bryan, T., McCutchan, T. F. and Chan, H. W. (1981) J. Virol. 39, 835-844
4. Noda, M., Kurihara, M. and Takano, T. (1982) Nucl. Acids Res. 10, 2865-2878
5. Mager, D. L. and Henthorn, P. S. (1984) Proc. Natl. Acad. Sci. USA 79, 4709-4713
6. Maeda, M., Otsuki, M., Yuu, H., Saeki, S., Yamasaki, T. and Bada, S. (1982) Eur. J. Cancer Clin. Oncol. 18, 123-128
7. Emi, M., Horii, A., Tomita, N., Nishide, T., Ogawa, M., Mori, T. and Matsubara, K. (1988) Gene 62, 229-235
8. Steel, P. E., Rabson, A. B., Bryan, T. and Martin, M. A. (1984) Science 225, 943-947
9. Nakamura, N. (1989) Med. J. Osaka Univ. 41, 121-132 (in Japanese)
10. Weiss, R., Teich, N., Varmus, H. and Coffins, S. (eds.) (1984) Cold Spring Harbor, New York
11. Rabson, A. B., Steele, P. E., Garon, C. F. and Martin, M. A. (1983) Nature (London) 306, 604-607
12. Rabson, A. B., Hamagishi, Y., Steele, P. E., Tykocinski, M. and Martin, M. A. (1985) J. Virol. 56, 176-182
13. Cohen, M., Powers, M., O'Connell, C. and Kato, N. Viology 147, 449-458
14. Kato, N., Larsson, E. and Cohen, M. (1988) Int. J. Cancer 41, 380-385
15. Cohen M., Kato, N. and Larsson, E. (1988) J. Cell. Biochem. 36, 121-128
16. Ono, M., Kawasaki, M. and Ushikubo, H. (1987) J. Virol. 61, 2059-2062
17. Gattoni-Celli, S., Kirsch, K., Kalled, S. and Isselbacher, K. J. (1986) Proc. Natl. Acad. Sci. USA 83, 6127-6131
18. Kato, N., P-Ohlsson, S., Kato, M., Larsson, E., Rydnert, J., Ohlsson, R. and Cohen, M. (1987) J. Virol. 61, 2182-2191
19. Izumi, T., Nagata, Y., Okabe, T., Sasaki, K., Kitamura, S., Kosaka, M., Kodama, T. and Yamaguchi, K. (1981) Nippon Kyobu Rinsho 40, 210-214 (in Japanese)
20. Chigwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979) Biochemistry 18, 5294-5299
21. Horii, A., Emi, M., Tomita, N., Nishide, T., Ogawa, M., Mori, T. and Matsubara, K. (1987) Gene 60, 57-64
22. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103-119
23. Feinberg, A. P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13
24. Southern E. M. (1975) J. Mol. Biol. 98, 503-517
25. Lehrach, H., Diamond, D., Wozney, J. M. and Boedtker, H. (1977) Biochemistry 16, 4743-4751
26. Gubler, U. and Hoffman, B. J. (1983) Gene 25, 263-269
27. Benton, W. D. and Davis, R. W. (1977) Science 196, 180-182
28. Sanger, F., Coulson, A. R., Barrell, B. G., Smith A. J. H. and Roe, B. A. (1980) J. Mol. Biol. 143, 161-178
29. Hattori, M. and Sakaki, Y. (1986) Anal. Biochem. 152, 232-238
30. Wilbur, W. J. and Lipman, D. J. (1983) Proc. Natl. Acad. Sci. USA 80, 726-730
31. Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedmann, T. and Schmid, C. W. (1981) J. Mol. Biol. 151, 17-33
32. Temin, H. M. (1982) Cell 28, 3-5
33. Yamamoto, T., Davis, C. G. Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L. and Russell, D. W. (1984) Cell 39, 27-38
34. Oikawa, S., Nakazato, H. and Kosaki, G. (1987) Biochem. Biophys. Res. Commun. 142, 511-518