Volume 212, number 2, 208–212 FEB 04431 February 1987

Cloning of Alu-containing cDNAs from human fibroblasts and identification of small Alu$^+$ poly(A)$^+$ RNAs in a variety of human normal and tumor cells

Svetlana A. Limborska, Sergei A. Korneev, Natalia E. Maleeva, Petr A. Slominsky, Appolon G. Jincharadze*, Pavel L. Ivanov* and Alexei P. Ryskov*

Institute of Molecular Genetics, USSR Academy of Sciences, Kurchatov Sq. 46, Moscow 123182 and **Institute of Molecular Biology, USSR Academy of Sciences, Vavilov str. 32, 117984, Moscow B-334, USSR*

Received 28 November 1986

Two clones have been selected from a human fibroblast cDNA bank. By DNA sequencing the clones were shown to contain Alu elements located near the ends of the cDNA inserts. DNA of the clones was used for Northern blot hybridization analysis of a number of poly(A)-containing RNAs from normal human tissues (brain, stomach, uterus, spleen, fibroblasts) and tumors (neurinoma, glioma, neuroblastoma, liposarcoma, adrenal cortex adenocarcinoma). All RNA samples reveal a heterodisperse distribution of Alu transcripts with discrete bands in the region of 7–12 S RNA. The majority of these small poly(A)$^+$ Alu$^+$ RNAs contain Alu sequences only in one (canonical) orientation with functional signals including the split promoter for RNA polymerase III.

Alu sequence; cDNA; Cloning; Transcription; (Human)

1. INTRODUCTION

In the human genome Alu DNA constitutes the dominant family of short repetitive sequences [1]. Each of these sequences is about 300 base pairs (bp) long, a head-to-tail dimer consisting of approx. 130 bp monomer units with an additional 31 bp segment located within the right monomer unit [2]. Alu sequences are efficiently transcribed, accounting for a significant fraction of hnRNA. In the course of processing, the majority of Alu transcripts are degraded. A small part survives and appears in the cytoplasm. Obviously, it is necessary to examine the transcription of repeated sequences in order to evaluate their possible functions. Recently the transcription properties of Alu-like elements (B1 and B2 repeat families) in various mouse cells have been characterized [3].

Here we describe partial sequences of two cloned human fibroblast cDNAs containing Alu. The Alu sequence is shown to be located near the 3′-end corresponding to poly(A)$^+$ RNA. We have also investigated major cytoplasmic poly(A)$^+$ transcripts containing Alu in different human tissues. We found an asymmetric representation of complementary strands of the Alu element in all cases studied. Discrete small Alu$^+$ poly(A)$^+$ RNAs (7–12 S) were detected in normal human tissues (brain, stomach, uterus, spleen, fibroblasts) and tumors (neurinoma, glioma, neuroblastoma, liposarcoma, adrenal cortex adenocarcinoma).

Correspondence address: S.A. Limborska, Institute of Molecular Genetics, USSR Academy of Sciences, Kurchatov Sq. 46, Moscow 123182, USSR

2. MATERIALS AND METHODS

cDNA clones were selected from a cDNA library prepared with cytoplasmic poly(A)$^+$ RNA of

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

human fibroblasts [4]. The DNA of clones was prepared as described in [5].

Plasmid inserts were sequenced by the method of Maxam and Gilbert [6]. Fragments to be sequenced were 3′-labeled at their *Pst*I ends with terminal deoxynucleotidyl transferase and [α-^{32}P]-dideoxyadenosine triphosphate. DNA strands were separated as in [7].

Polyadenylated RNAs were isolated as described in [8]. All the RNA preparations were purified by two precipitations with 3 M sodium acetate. Aliquots (5 μg) of RNA were fractionated by electrophoresis in a 1% agarose gel containing 6% formaldehyde. The RNA was transferred to a nitrocellulose filter and hybridized with ^{32}P-labeled DNA strands as in [7].

3. RESULTS

From the cDNA library prepared with cytoplasmic poly(A)$^+$ RNA of human fibroblasts, we selected several clones hybridizing with total human DNA and thus representing repetitive elements in the cDNAs. Three such clones were purified and grown. DNA was isolated, treated with restriction enzymes *Eco*RI or *Pst*I and analysed by the Southern method using ^{32}P-labeled total human DNA for hybridization. The results presented in fig.1 indicate that the two cDNA insertion hybridizing with total DNA (pRHF11 and pRHF40) are about 400 bp long. The *Pst*I site of pRHF24 is probably lost and its insertion could not be cut out.

Fig.2 shows the nucleotide sequences of the cDNAs from pRHF11 and pRHF40. Both contain an Alu repetitive sequence, which seems to be located near the 3′-end of the corresponding RNA transcript. A full-length copy of Alu (287 bp) was found in pRHF40 and only a part (148 bp) in pRHF11. The latter result is probably due to the loss of the right Alu monomer during cDNA synthesis and cloning. Both Alu sequences are typical members of the Alu family. They show 12 and 14% divergence from the human Alu consensus sequence [9]. This accords with the average divergence (13%) within the human Alu family.

Computer analysis of the pRHF11 sequence shows that one of the direct translation frames contains no stop codons. In the other two frames, stop codons are found in the middle of the sequence. In all cases ORFs are spread over the Alu zone. Similar analysis of the pRHF40 sequence shows the longest ORF (about 250 bp) in one of the direct translation frames. This ORF extends over half of an Alu sequence corresponding to the size of the Alu zone of clone pRHF11. Alu elements in pRHF40 and pRHF11 do not contain the AATAAA polyadenylation signal, which is a guide for detecting post-transcriptional poly(A) and the 3′-end of RNA. Therefore, the exact orientation and location of the Alu sequences in cloned RNA transcripts remain unknown. Nevertheless, our results demonstrate that the hybridization of the Alu sequences with cytoplasmic poly(A)$^+$ RNA described by others [10] may depend on the presence of typical Alu elements in the RNA.

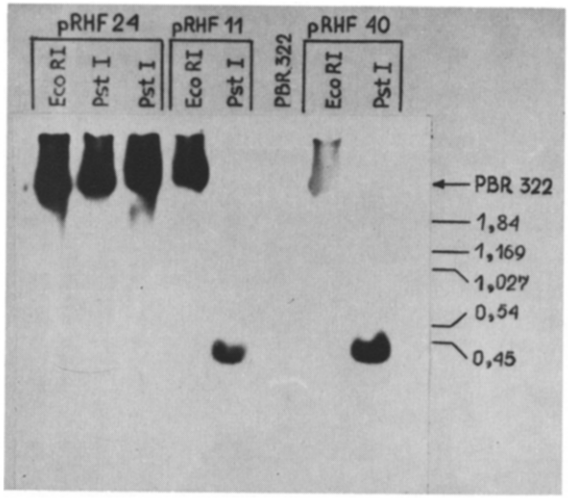

Fig.1. Southern blot analysis of plasmid DNAs containing repetitive sequences. Total human DNA was labeled by nick translation and used as a probe for hybridization with plasmid DNAs digested with *Eco*RI (lanes 1,4,6,7) or *Pst*I (lanes 2,3,5,8). *Hin*dIII fragments of SV40 DNA were used as markers.

DNA inserts in pRHF11 and pRHF40 or their separate strands were further used for hybridization analysis in order to examine the distribution of Alu sequences among mRNAs in different human tissues (fig.3) and to establish the orientation of Alu in these RNAs (fig.4). We designated a strand containing signal sequences (like the RNA polymerase III promoter) as a 'plus (+) strand' of 'canonical strand' (see fig.2) and the complemen-

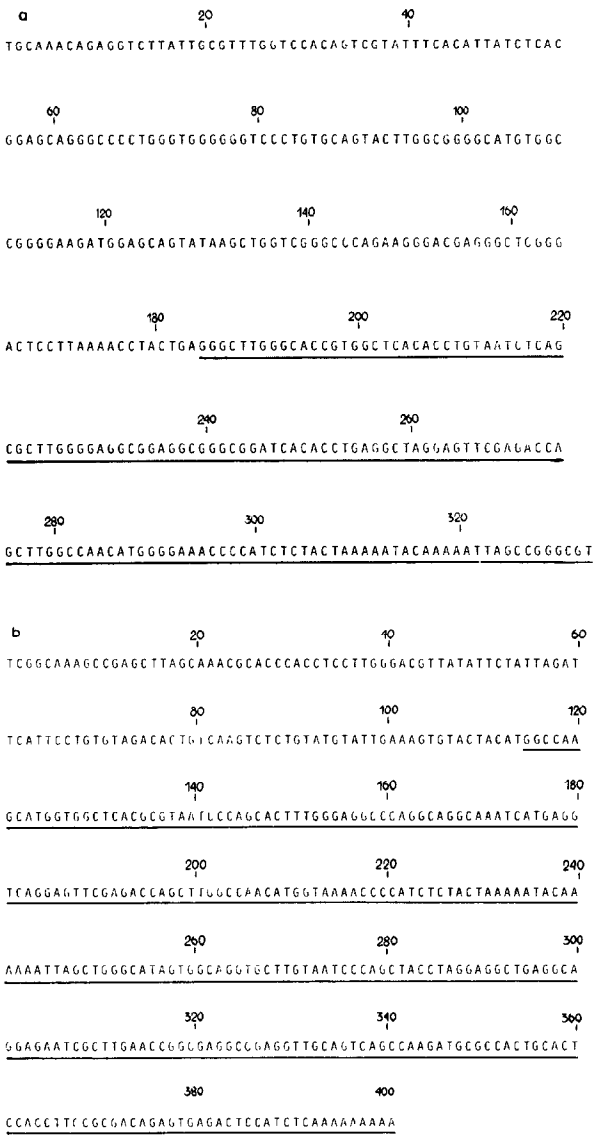

Fig.2. Nucleotide sequences of human DNA inserts in pRHF11 (a) and pRHF40 (b). Plus (+) strands corresponding to those present in poly(A)$^+$ RNA are shown in the 5′ to 3′ direction. Alu sequences are underlined.

tary strand as a 'minus (−) strand'. We observed (fig.3) a heterogeneous distribution of Alu sequences in all the cytoplasmic poly(A)$^+$ RNAs studied (from spleen, adrenal cortex adenocarcinoma, stomach, neurinoma, brain). A prominent hybridization band was observed in the zone of 7 S RNA for every sample except spleen RNA, where

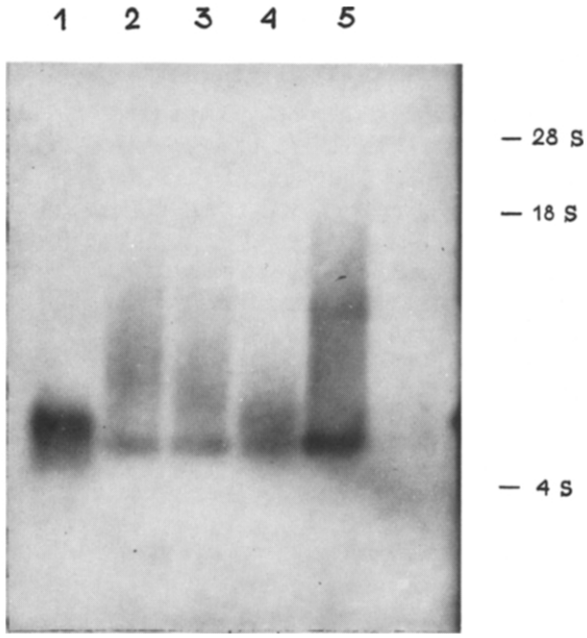

Fig.3. Northern blot hybridization of Alu-containing DNA insert from clone pRHF11 with poly(A)$^+$ RNA from several types of human tissues. Lanes: 1, cytoplasmic poly(A)$^+$ RNA from spleen; 2, adrenal cortex adenocarcinoma; 3, stomach; 4, neurinoma; 5, brain.

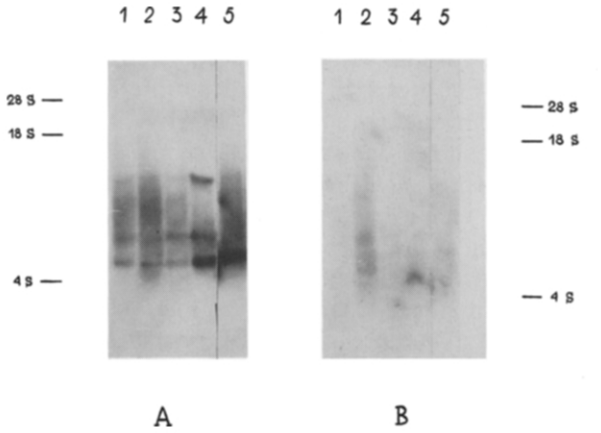

Fig.4. Northern blot hybridization of separated cDNA strands from clone pRHF11 with poly(A)$^+$ RNA from several types of human tissues. (A) Hybridization with the minus chain (−). (B) Hybridization with the plus chain (+). Lanes: 1, cytoplasmic poly(A)$^+$ RNA from neuroblastoma; 2, uterus; 3, fibroblasts; 4, stomach; 5, brain.

the main hybridization material was situated in a higher-M_r RNA fraction. In the brain, one can see two hybridization bands: at 7 S and 12 S RNAs. In a situation of higher hybridization efficiency (e.g. higher specific radioactivity of labeled DNA) it is possible to detect additional bands of Alu$^+$ RNA (fig.4). Hybridization experiments with separated strands of the Alu element (fig.4) showed cytoplasmic poly(A)$^+$ RNA to be bound mainly to the (−) strand: it contained almost exclusively Alu sequences in the canonical (+) orientation. Three distinct zones of hybridization with the (−) strand of Alu were visible in all tissues. The same results were obtained with RNA from glioma and liposarcoma (not shown). Hybridization bands are most prominent in the brain. Similar hybridization patterns were obtained with the clones pRHF11 and pRHF40 or their parts containing only the Alu sequence. This means that all hybridization RNAs contain at least a part of Alu in the canonical orientation. To exclude the possibility that low-M_r Alu$^+$ RNA was not polyadenylated but bound to another polyadenylated RNA, the following control experiment was carried out. Poly(A)$^+$ RNA from human adrenal cortex adenocarcinoma was denatured in 100% formamide at 75°C, diluted 100-fold with cold buffer (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.2% SDS) and repeatedly chromatographed on poly(U)-Sepharose. Northern filters with RNA samples thus fractionated were hybridized with Alu DNA. One can see (fig.5) that only the material repeatedly bound on poly(U)-Sepharose contained small RNAs hybridizing with Alu DNA. Therefore, we can conclude that these low-M_r Alu$^+$ RNAs are polyadenylated.

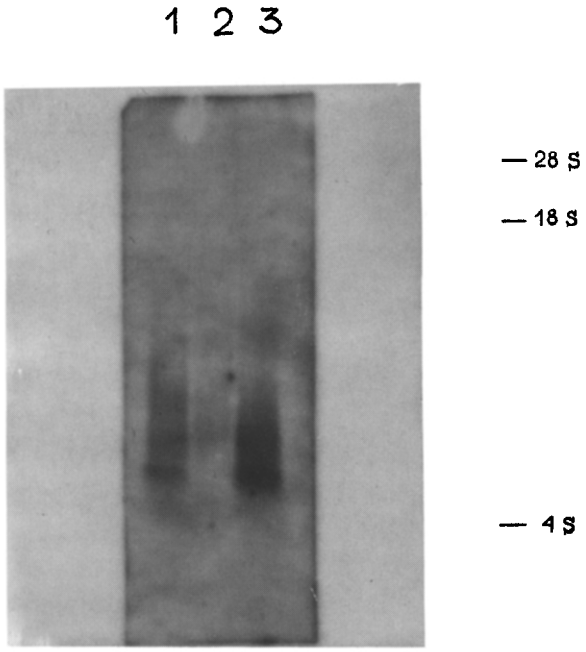

Fig.5. Blot hybridization analysis of Alu-containing RNA after melting and subsequent rechromatography on poly(U)-Sepharose. Hybridization of the cDNA insert from clone pRHF11 was performed with poly(A)$^+$ RNA from adrenal cortex adenocarcinoma. Lanes: 1, hybridization with poly(A)$^+$ RNA before melting; 2,3, after melting and rechromatography on poly(U)-Sepharose; hybridization with nonadsorbed RNA (2); hybridization with adsorbed RNA (3).

4. DISCUSSION

Previous studies have demonstrated in vivo and in vitro transcription of human Alu and different rodent family members [3,10–12]. A portion of the RNA component of the signal recognition particle (7 S L RNA) shares homology with the Alu sequence [13]. The presence of Alu elements in the intervening sequences and untranslated regions of a number of structural genes also contributes to the observed expression of the Alu family [14–16]. Jelinek et al. [2] and Tashima et al. [17] have shown a high concentration of Alu sequences in nuclear RNA and a much lower concentration in total cytoplasmic poly(A)$^+$ RNA. However, among the heterogeneous distribution of Alu$^+$ RNA, the authors did not detect any discrete fraction of small Alu$^+$ RNA.

Here, we have demonstrated the existence of distinct small Alu$^+$ poly(A)$^+$ RNAs in several human tissues. Like 7 S L RNA, these transcripts are homologous to the canonical (+) strand of Alu, but, in contrast to the 7 S L RNA, they are polyadenylated.

The origin of small Alu$^+$ poly(A)$^+$ RNAs is not clear. They can be primary transcripts of Alu elements which are synthesized with the aid of RNA polymerase III. For example, this way was proved for small mouse B2$^+$ poly(A)$^+$ RNAs [18]. Alternatively, they can be transcribed by RNA polymerase II as parts of long transcripts from which they are later excised in the course of pro-

cessing. These small RNAs may be involved in the regulation of gene expression or the retroposition of Alu sequences to new sites in the genome.

ACKNOWLEDGEMENT

We thank Professor G.P. Georgiev for a helpful discussion.

REFERENCES

[1] Houck, C.M., Reinhart, F.P. and Schmid, C.W. (1979) J. Mol. Biol. 132, 289–306.
[2] Jelinek, W.R., Toomey, T.P., Leinwand, L., Duncan, C.H., Biro, P.A., Choudary, P.V., Weissman, S.M., Rubin, C.M., Honk, C.M., Deininger, P.L. and Schmid, C.W. (1980) Proc. Natl. Acad. Sci. USA 77, 1398–1402.
[3] Ryskov, A.P., Ivanov, P.L., Tokarskaya, O.N., Kramerov, D.A., Grigoryan, M.S. and Georgiev, G.P. (1985) FEBS Lett. 182, 73–76.
[4] Ryskov, A.P., Korneev, S.A., Prosnjak, M.I. and Limborska, S.A. (1983) Dokl. Akad. Nauk SSSR 272, 1268–1272.
[5] Birnboim, H.C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513–1518.
[6] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–508.
[7] Ryskov, A.P., Ivanov, P.L., Kramerov, D.A. and Georgiev, G.P. (1983) Nucleic Acids Res. 11, 6541–6558.
[8] Schibler, U., Tosi, M., Pittet, A.C., Fabiani, L. and Wellauer, P.K. (1980) J. Mol. Biol. 142, 93–116.
[9] Deininger, P.L., Jolly, D.J., Rubin, C.M., Friedmann, T. and Schmid, C.W. (1981) J. Mol. Biol. 151, 17–33.
[10] Elder, J.T., Pan, J., Duncan, C.H. and Weissman, S.M. (1981) Nucleic Acids Res. 9, 1171–1189.
[11] Jackson, M., Heller, D. and Leinwand, L. (1985) Nucleic Acids Res. 13, 3389–3403.
[12] Perlino, E., Paonessa, G. and Ciliberto, G. (1985) Nucleic Acids Res. 13, 8359–8377.
[13] Ullu, E. and Tschudi, C. (1984) Nature 312, 171–172.
[14] Tsukada, T., Watanabe, Y., Nakai, Y., Imura, H., Nakanishi, S. and Numa, S. (1982) Nucleic Acids Res. 10, 1471–1478.
[15] Yamamoto, T., Davis, C.G., Brown, M.S., Schneider, W.J., Casey, M.L., Goldstein, J.L. and Russel, D.W. (1984) Cell 39, 27–38.
[16] Lee, M.G., Loomis, C. and Cowan, N.J. (1984) Nucleic Acids Res. 14, 5823–5836.
[17] Tashima, M., Calabretta, B., Torelli, G., Scofield, M., Maisel, A. and Saunders, G.F. (1981) Proc. Natl. Acad. Sci. USA 78, 1508–1512.
[18] Kramerov, D.A., Tillib, S.V., Lekakh, I.V., Ryskov, A.P. and Georgiev, G.P. (1985) Biochim. Biophys. Acta 824, 85–98.