Molecular and Cellular Biology, July 1993, p. 4233–4241
0270-7306/93/074233-09$02.00/0
Copyright © 1993, American Society for Microbiology
Vol. 13, No. 7

Multiple Dispersed Loci Produce Small Cytoplasmic *Alu* RNA

RICHARD J. MARAIA,[1]* CLAIRE T. DRISCOLL,[1] TIMOTHY BILYEU,[2] KARL HSU,[1] AND GRETCHEN J. DARLINGTON[2]

Laboratory of Molecular Growth Regulation, National Institute of Child Health and Human Development, Bethesda, Maryland 20892,[1] and Department of Pathology, Texas Children's Hospital, Houston, Texas 77030[2]

Received 12 January 1993/Returned for modification 22 February 1993/Accepted 15 April 1993

Alu* repeats are short interspersed elements (SINEs) of dimeric structure whose transposition sometimes leads to heritable disorders in humans. Human cells contain a poly(A)$^-$ small cytoplasmic transcript of ~120 nucleotides (nt) homologous to the left *Alu* monomer. Although its monomeric size indicates that small cytoplasmic *Alu* (sc*Alu*) RNA is not an intermediary of human *Alu* transpositions, a less abundant poly(A)-containing *Alu* transcript of dimeric size and specificity expected of a transposition intermediary is also detectable in HeLa cells (A. G. Matera, U. Hellmann, M. F. Hintz, and C. W. Schmid, Mol. Cell. Biol. 10:5424–5432, 1990). Although its function is unknown, the accumulation of *Alu* RNA and its ability to interact with a conserved protein suggest a role in cell biology (D.-Y. Chang and R. J. Maraia, J. Biol. Chem. 268:6423–28, 1993). The relationship between the ~120- and ~300-nt *Alu* transcripts had not been determined. However, a B1 SINE produces scB1 RNA by posttranscriptional processing, suggesting a similar pathway for sc*Alu*. An *Alu* SINE which recently transposed into the neurofibromatosis 1 locus was expressed in microinjected frog oocytes. This neurofibromatosis 1 *Alu* produced a primary transcript followed by the appearance of the sc*Alu* species. 3′ processing of a synthetic ~300-nt *Alu* RNA by HeLa nuclear extract in vitro also produced sc*Alu* RNA. Primer extension of sc*Alu* RNA indicates synthesis by RNA polymerase III. HeLa-derived sc*Alu* cDNAs were cloned so as to preserve their 5′-terminal sequences and were found to correspond to polymerase III transcripts of the left monomeric components of three previously identified *Alu* SINE subfamilies. Rodent × human somatic cell hybrids express *Alu* RNAs whose size, heterogeneous length, and chromosomal distribution indicate their derivation from SINEs. The coexpression of dimeric and monomeric *Alu* RNA in several hybrids suggests a precursor-product relationship.

Mouse B1 and human *Alu* repeats are species-specific members of the *Alu* family of short interspersed elements (SINEs) and share approximately 80% sequence homology (16, 45). *Alu* repeats are transposed sequences present at approximately 5×10^5 copies per haploid genome in human DNA. Although a function for *Alu* SINEs is unknown, their repetitive nature and propensity for mobility have led to genetic variability and heritable disorders in humans by both DNA- and RNA-mediated events (31, 51). Structural evidence indicates that *Alu* SINEs are retrotransposed through an intermediary transcript synthesized by RNA polymerase III (Pol III) (41, 53). Recent evidence suggests that Pol III-synthesized *Alu* RNAs accumulate in vivo (7, 30).

The efforts of several research groups revealed that B1 and *Alu* sequences can each be classified into subfamilies of different evolutionary ages (4, 18, 34, 35, 48, 54). Although the genes that gave rise to the subfamilies have not been identified, their sequences have been reconstructed and represented by subfamily consensuses. Accumulated mutations render older elements more divergent than younger ones in nucleotide sequence (2, 17, 29, 34). Analyses of recently transposed *Alu*s revealed that they belong to a discernibly young subfamily (2, 29) which contains only ~0.2% of the total number of human *Alu* repeats (30). Young *Alu* subfamilies contain transpositionally active elements, while more divergent *Alu*s may be relatively idle (30, 47). Thus, *Alu* sequences continue to expand and evolve within the human genome. *Alu* transpositions such as recently occurred in the neurofibromatosis 1 (NF1) gene (51) appear to be derived from a select subset of *Alu*s (4, 44), comprised of a small number of active progenitors (2, 17, 23, 29, 44, 54), which exhibit a high propensity for transposition. Once an *Alu* source gene is identified, the study of its transcripts and patterns of expression should aid our understanding of these elements and their mobility. At present, little is known about Pol III-derived *Alu* transcripts expressed in vivo.

Although several observations indicate that B1 and *Alu* SINEs are homologs (see below), these SINEs differ in primary structural organization. A typical young *Alu* comprises a ~280-bp imperfect dimer, while B1 is an ~135-bp monomer. Although a few monomeric *Alu*s have been identified in human DNA (19, 34, 36, 37), they comprise highly divergent sequences, suggesting that they are vestiges of early *Alu* evolution derived from an ancient monomeric *Alu*-like sequence that was transpositionally active in an ancestral primate genome. A monomeric, poly(A)-containing *Alu*-related transcript of 200 nucleotides (nt) named BC200 is detectable in Old World monkey (52) and human (50) brain. Tiedge and coworkers (50) suggest that a left *Alu* monomer evolved into the BC200 gene, which now gives rise to the 200-nt polyadenylated *Alu* RNA. Apparently, *Alu* monomeric elements developed into the more proliferative dimeric *Alu* SINEs presently found in the great apes and humans. While dimeric *Alu*s have continued to evolve and expand along their own pathway, the transpositional activity, if any, of monomeric *Alu*s appears to be relatively idle in higher primates (for discussions, see references 8, 17, 19, and 37).

Thus, rodent B1 repeats exist in their genome as transposed monomers, whereas active human *Alu*s are transposed as dimers. Despite their difference in primary organization, B1 and *Alu* repeats share several notable characteristics. A young B1 contains an internal split pro-

* Corresponding author.

moter that directs transcription by Pol III and is followed by a ~40-nt poly(A) tract (35, 55). The ~120-bp left *Alu* monomer also contains a promoter homologous to B1, while the right monomer instead contains a 31-bp stretch not found in the left monomer (10, 12). The two *Alu* monomers are connected by ~20-bp of A-rich spacer, while the right monomer is followed by a ~40-bp poly(A) tract (2, 29, 51). As indicated above, the B1 and *Alu* sequences have evolved through similar successions of founder or source genes and share structural features characteristic of SINE retrotransposons. The notion that both use an intermediary transcript in similar mechanisms of retrotransposition is supported by the evolutionary conservation of a complex *Alu* domain RNA secondary structure by successive B1 and *Alu* subfamily consensus sequences as well as transcripts of newly inserted B1 and *Alu* repeats (despite significant sequence difference) but not random B1 and *Alu* elements (17, 20, 21, 46).

A low-abundance *Alu* RNA of dimeric size (~300 nt) and sequence specificity expected of a putative transposition intermediary was identified in human tissue culture cells (30). However, although the great majority of transposed human *Alu*s represent dimeric DNA elements, recent investigations using highly specific oligonucleotide DNA probes as well as sensitive Northern (RNA) analysis indicate that a cellular *Alu* RNA of ~120 nt accumulates in the cytoplasm of HeLa cells as a monomeric species that lacks poly(A) (7, 30) (see Fig. 3). This poly(A)$^-$ *Alu* RNA species is also present in several transformed cell lines as well as nontransformed human cells and normal human tissues (8a), suggesting a role in cell biology. This species is distinguishable from the previously described BC200 RNA by several criteria [e.g., size, sequence, and lack of poly(A)] and will hereafter be referred to as small cytoplasmic *Alu* (sc*Alu*) RNA. Interestingly, *Alu* monomeric transcripts, like their rodent counterparts scB1 transcripts (26), also display the aforementioned RNA structure (46; this report). On the basis of their relatedness of primary sequence, secondary structure, abundance, cytoplasmic partitioning, and ability to specifically and interchangeably interact in vitro with a conserved protein present in both rodent and human cells, it was suggested that human sc*Alu* and mouse scB1 RNAs may represent functional homologs even though no function has yet been ascribed to them (7). Thus, the *Alu* domain RNA structure which was conserved by the *Alu* and B1 sequences that generate new transpositions is also found in the small cytoplasmic transcripts that accumulate in vivo. These observations suggest that *Alu* and B1 families consist of RNA-encoding genes that evolved under selective pressure plus a large number of transposed copies of these genes. At present, the exact relationship between sc*Alu* RNA, the larger dimeric *Alu* RNA, *Alu* SINEs, and the source genes of de novo *Alu* transpositions is unknown.

It was suggested that sc*Alu* RNA is produced by 3′ posttranscriptional processing of *Alu* primary transcripts in a manner previously described for its mouse homolog B1 (30). In that case, a recently transposed B1 element (35) is transcribed by Pol III into a poly(A)-containing primary transcript that subsequently undergoes 3′ processing to produce a poly(A)$^-$ small stable cytoplasmic RNA (1, 26–28). By analogy, sc*Alu* RNA may be derived from one or more *Alu* SINEs or possibly from an *Alu* source gene(s). It would be of interest therefore to further characterize sc*Alu* RNA by examining its complexity and determine whether the sequences which encode it are recognizably different in structure and or distribution from typical *Alu* SINEs.

In this study, we characterized sc*Alu* RNA by primer extension, cDNA sequence, and chromosomal origin. Analysis of multiple sc*Alu* cDNAs revealed sequences specific to the left monomeric components of three previously identified dimeric *Alu* SINE subfamilies, including the youngest subfamily, PV (predicted variant). These sequences displayed a significantly higher degree of preservation of the predicted *Alu* domain RNA secondary structure than did random *Alu*s. We also demonstrated that some rodent × human somatic cell hybrids express *Alu* RNA. Furthermore, cloned human-specific Alu SINEs microinjected into frog oocyte nuclei produced primary transcripts and monomeric *Alu* RNA with kinetics suggestive of posttranscriptional processing. Finally, in vitro-synthesized dimeric *Alu* SINE RNA subjected to HeLa nuclear extract demonstrated 3′ processing and generation of sc*Alu* RNA. Cumulatively, the data show that the sc*Alu* RNA species can be generated from typical dimeric *Alu* SINEs by RNA processing and that multiple loci dispersed on several different human chromosomes encode this RNA.

MATERIALS AND METHODS

Hybrid cell lines were from two sources. O. Pereira-Smith provided lines A9+2, A9+3, A9+12, A9+15, which have been characterized previously as human monochromosomal (33). Cell lines GM 07299, 10478, 10479, 10498, 10567, 11010, 10449, 10498, 10888, 10114, 10115, 10629, 10898, 10791, 06317, 10611, 10516B, and 10926B were obtained from the Coriell Institute for Medical Research (Camden, N.J.) and have also been extensively characterized as described in the 1992–1993 *Catalog of Cell Lines* (32a). These cell lines were maintained in selective media as recommended. Our verification of the human chromosomes in hybrid clones was done by Giemsa-11 staining, which differentially stains rodent and human chromosomes (3). At least 15 metaphases were analyzed for each hybrid clone. The majority of clones contained the expected single human chromosome(s) at a frequency within 10% of that published.

Total RNA was prepared as previously described (26). All cells were grown as monolayers. The attached cells were washed with phosphate-buffered saline (PBS) three times before the addition of guanidinium lysate buffer (4 M guanidinium thiocyanate, 10 mM Tris-Cl [pH 7.4], 0.5% *n*-lauroyl sarcosine, 1% β-mercaptoethanol). The samples were homogenized with a Polytron to shear genomic DNA and made 0.2 M sodium acetate (NaAc) (pH 4.0). The samples were extracted with H_2O-saturated (acidic) phenol-chloroform twice and ethanol precipitated overnight. Total nucleic acid was resuspended in 0.5 ml of 10 mM Tris-Cl (pH 7.5)–10 mM $MgCl_2$–5 mM VRC (Bethesda Research Laboratories) and treated with 8 μl of DNase I (RNase free; Boehringer Mannheim) at 37°C for 15 min. Samples were then made 1% sodium dodecyl sulfate–0.15 M NaAc (pH 4.0), extracted twice with acidic phenol-chloroform as described above, chloroform extracted, and ethanol precipitated. The integrity of the RNA samples was verified by polyacrylamide gel electrophoresis (PAGE) and ethidium staining, which revealed a characteristic profile of small cellular RNAs with no evidence of degradation. Control experiments using in vitro-synthesized [^{32}P]RNA added to whole cell lysates prior to RNA purification revealed that full-length *Alu* transcripts remain intact and do not generate sc*Alu* RNA by artifactual processing (data not shown).

RNA blotting was done from 5.5% polyacrylamide (40:1 acrylamide/bisacrylamide)–8 M urea gels as previously de-

scribed (26). One hundred micrograms of cellular RNA (as an ethanol suspension) plus 40 μg of sonicated salmon sperm carrier DNA (RNase free; ethanol suspension), the latter of which facilities transfer (13), were copelleted, dried, and resuspended in formamide loading buffer. This material was fractionated by PAGE and transferred to GeneScreen Plus (Dupont) in a Bio-Rad trans-blot cell with plate electrodes at 30 V for 16 h with continuous cooling. The blot was cross-linked with a Stratagene UV cross-linker and allowed to dry thoroughly before prehybridization and hybridization were done.

For cytoplasmic RNA preparation, HeLa cells were washed three times with PBS and scraped from the monolayer. The cells were washed two additional times with ice-cold PBS, each time collecting cells by centrifugation. PBS was drained, and cells were resuspended in low-salt lysis buffer (10 mM Tris-Cl [pH 7.5], 10 mM $MgCl_2$, 5 mM VRC, 0.5% Nonidet P-40) and put on ice for 10 min. Cells were then subjected to 20 strokes with a Dounce homogenizer (pestle B), and nuclei were collected by centrifugation. The cytoplasmic supernatant was made 0.2 M NaAc (pH 4.0) and extracted with acidic phenol-chloroform as described above.

Hybridization to a probe (designated Alu-24) complementary to bases 65 to 88 of the precise and PV subfamilies was previously described (7).

sc*Alu* cDNA was prepared by primer extension using the ^{32}P-end-labeled oligonucleotide DNA 5′-AGT AGA GAC GGG GTT TCA CC-3′ as described previously (26). HeLa cell RNA was fractionated by high-resolution 6% polyacrylamide–8 M urea gel electrophoresis. A band encompassing the ~110- to 125-nt size range was excised, and the RNA was eluted. After ethanol precipitation, this material was used for cDNA synthesis. An 118-nt sc*Alu* RNA was synthesized in vitro from the left monomer of the NF1 *Alu* (51), using a T7 RNA polymerase-dependent template constructed by the polymerase chain reaction (PCR) (26); 0.01 pg of this RNA was used for positive control primer extension reactions. Poly(dA) tailing, PCR amplification, and *Hin*dIII-*Sal*I directional cloning into pUC18 were done as described previously (26) except that the 3′ PCR primer was 5′-GTG CAA TTA AGC TTA GTA GAG ACG GGG TTT C. Chain termination sequencing was performed.

Microinjection of the germinal vesicles of mature oocytes with NF1 *Alu*-containing plasmid DNA (51) and [α-^{32}P]GTP was done as described previously (26). At various times thereafter, RNA was purified and separated by PAGE. Nonradioactive, oocyte-injected RNA was used for oligo(dT) chromatography. Poly(A)$^+$ and poly(A)$^-$ RNAs were separated by oligo(dT) spin columns (Pharmacia).

In vitro RNA processing substrate was generated from the *Alu* which resides in the fourth intron of the human alpha-fetoprotein gene (38). RNA was synthesized from a T7 RNA polymerase-dependent template constructed by PCR to produce an RNA whose 5′ end is coincident with initiation by Pol III. After dephosphorylation, the RNA was labeled on its 5′ end with [γ-^{32}P]ATP and polynucleotide kinase (26). The product was purified after denaturing PAGE and subjected to HeLa cell nuclear extract to which RNasin (Promega) was added to 4 U/μl as described previously (26). At various times thereafter, equal aliquots were removed and reaction were stopped; then RNA was purified and analyzed by denaturing PAGE and autoradiography.

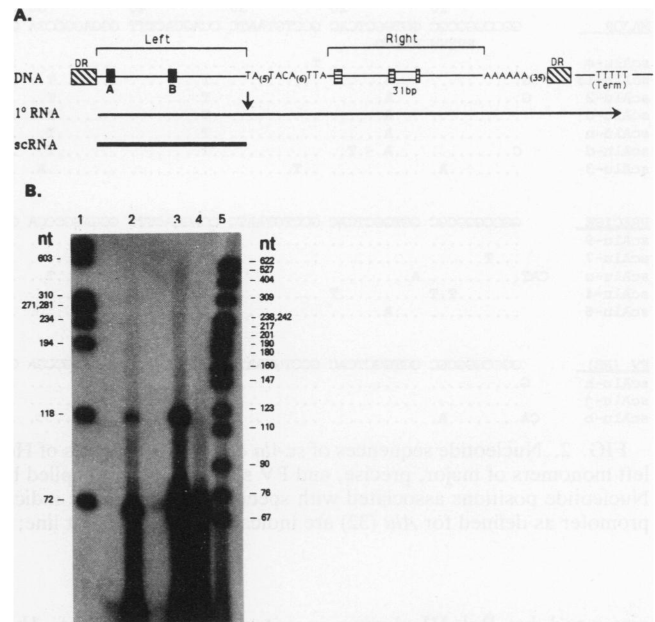

FIG. 1. (A) Relationship of a transposed *Alu* and its Pol III-dependent transcripts. Human *Alu* SINEs are dimers that are flanked by direct repeats (DR). The left monomer contains a promoter consisting of A and B boxes (filled) that directs Pol III to initiate transcription at or near the first base of the repeat (10, 12). A potential B box of the right monomer is interrupted by a 31-bp sequence (open rectangle) (10, 12) not found in the left monomer. Transcriptional termination may occur at the first $(dT)_{\geq 4}$ tract beyond the 3′ DR to produce a primary (1°) RNA. Posttranscriptional processing may convert the primary transcript to sc*Alu* RNA whose 3′ end is depicted by the downward arrow. (B) Primer extension of sc*Alu* RNA. Size-fractionated sc*Alu* RNA from HeLa cells plus tRNA carrier was subjected to primer extension analysis using a ^{32}P-end-labeled primer complementary to the 3′ end of the left *Alu* monomer. An autoradiogram of the products following PAGE is shown. Lanes: 1, *Hae*III-digested ϕX174 ^{32}P-labeled size markers; 2, HeLa sc*Alu* RNA plus tRNA carrier; 3, *Alu* left-monomer RNA (0.01 pg) synthesized in vitro plus tRNA; 4, tRNA alone; 5, *Msp*I-digested pBR322 size markers.

RESULTS

sc*Alu* RNA corresponds to an RNA Pol III-initiated transcript. Although the previously characterized sc*Alu* RNA was demonstrated to correspond to a transcript of the left monomer by virtue of its size and reactivity with a left-monomer-specific probe (7, 30), the sequence and 5′ start site of this RNA had not been determined. To this end, HeLa cell sc*Alu* RNA was size fractionated and purified after denaturing PAGE and then analyzed by primer extension along with *Escherichia coli* tRNA carrier. A ^{32}P-labeled oligodeoxynucleotide 100% complementary to the 3′ end of the *Alu* left monomer was used; human subfamily consensus sequences are invariant in this region (30, 44). This procedure yielded a cDNA band that comigrated with the 118-nt marker, the size expected from an *Alu* RNA initiated by Pol III (Fig. 1B). This product was obtained from either HeLa scRNA (lane 2) or an *Alu* left-monomer RNA synthesized in vitro and used as a positive control (7, 26) (lane 3) but not carrier tRNA alone (lane 4). This novel finding provides evidence that sc*Alu* RNA is a specific product of Pol III transcription, since its 5′ end mapped to the known initiation

```
                  10         20         30         40         50         60           70         80         90        100        110      118
MAJOR        GGCCGGGCGC GGTGGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGCCGA GGCGGGCGGA TCACCTGAGGTC AGGAGTTCGA GACCAGCCTG GCCAACATGG TGAAACCCCG TCTCTACT
                  RRYNN RRYGG                                                                   GSTCRA NNCC
scAlu-e      .......... .......... T......... .......... .......... ......T... ...T........ .......... .......... ...G..G... .......... ........
scAlu-13    G.......... ...A...... .......... .T........ .......T.. ......T... .........A.. .......A.. .......... .......... .......... ........
scAlu-2     G.......... ...A...... .......... .T........ .......T.. ......T... .........A.. .......A.. .......... .......... .......... ........
scAlu-o      .......... ...A...... .......... .T........ .......T.. ......T... .........A.. .......A.. .......... .......... .......... ........
scAlu-n      .......... ...A...... .......... .T........ .......T.. ......T... .........A.. .......A.. .......... .......... .......... ........
scAlu-d     C.......... ...A.-.T.. .......... .T........ .......T.. ......T... .........A.. .......A.. .......... .....T.... .......... ........
scAlu-3      ........A. .......... ..T....... .......... ........A. ..AA..T.A. ....TA...... .......T.. .......... .......... .......... ........

                                                                                  **          *          *        *   *
PRECISE      GGCCGGGCGC GGTGGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGCCGA GGCGGGCGGA TCAC--GAGGTC AGGAGATCGA GACCATCCTG GCTAACACGG TGAAACCCCG TCTCTACT
scAlu-9      .......... .......... .........T .......... .......... .......... ....--...... ..C....... .......... .......... .......... ........
scAlu-7      ...T...... .......... .......... .......... .......... .......... ....--...A.. .A........ .....C.... ..C...T... .......... ........
scAlu-c   CAT.......... A......... .......... .......A.. .......T.. ....A..A.. ....--...... .......A.. .......... .......... .......... ........
scAlu-4      .......T.T .........T .......... .T........ .......... ...A...A.. ...T--...... .......T.. .......... ..C...T... .......... ........
scAlu-6      .......... ...A...... .......... .T........ .......... ......T... ..G.--...... ......T.A. .....G.... ..C..T.T.. .......... ........

                                                                                                                  *      *
PV (HS)      GGCCGGGCGC GGTGGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGCCGA GGCGGGCGGA TCAC--GAGGTC AGGAGATCGA GACCATCCCG GCTAAAACGG TGAAACCCCG TCTCTACT
scAlu-h     G.......... .......... .......... .......... .......... .......... ....--...... .......... .......... .......... .......... ........
scAlu-j      .......... .......... .......... .......... .......... .......... ....--...... .......... .......... .......... .......... ........
scAlu-b    CA........A. .......... .......... .......... .......... .......... ....--...... .......... .......... .......... .......... ........
```

FIG. 2. Nucleotide sequences of sc*Alu* cDNAs. Sequences of HeLa-derived sc*Alu* cDNAs are compared with consensus sequences of the left monomers of major, precise, and PV subfamilies as compiled by Matera et al. (29). The primer used for cDNA synthesis is underlined. Nucleotide positions associated with specific subfamilies are indicated by asterisks. The A- and B-box consensus sequences of the Pol III promoter as defined for *Alu* (32) are indicated below the first line; R = G or A; Y = C or T; S = T or A.

site used by Pol III during in vitro transcription of *Alu* elements (10, 12).

The resulting full-length sc*Alu* cDNA was poly(dA) tailed by using terminal deoxynucleotide transferase, PCR amplified, and cloned into pUC vector as previously described for scB1 RNA (26). This method uses a poly(dT) primer (11) instead of an *Alu* sequence for PCR amplification, thereby decreasing the likelihood of inadvertent amplification of *Alu* sequences interspersed in high-molecular-weight nucleic acids and preserving the 5′-most sequence information. The presence of a poly(dT) tract immediately 5′ to each of the sc*Alu* cDNA sequences indicates derivation from Pol III-initiated *Alu* RNA (data not shown). The sequences of 15 sc*Alu* cDNAs were determined and are shown in Fig. 2. The presence of additional bases 5′ to the *Alu* consensus in some of the cDNAs (clones 13, 2, c, h, and b) was not surprising since it was previously demonstrated that transcription initiation by Pol III may begin at or just before the first base of a given repeat (1, 10, 12). We must note that microheterogeneity may also exist at the 3′ termini of sc*Alu* RNAs, since the method used precluded ascertainment of this information. Thus, HeLa sc*Alu* RNA exhibits sequence and length microheterogeneity. High-resolution PAGE-Northern analysis in which HeLa RNA was coelectrophoresed and blotted along with RNA size markers (data not shown) indicated that the sizes of sc*Alu* RNAs ranged from 115 to 120 nt, in good agreement with the sequence data of Fig. 2.

Several groups classified *Alu* repeats into subfamilies of different evolutionary ages (4, 18, 34, 48, 54). Although their consensus sequences agree, a universal subfamily nomenclature has not yet been established. For the purposes of this report, we adopted major, precise, and predicted variant (PV) subfamily designations as have been used in other recent publications (23, 39). The sc*Alu* cDNAs represent the left monomeric components of all three of these *Alu* subfamilies. Figure 2 shows that sc*Alu* cDNA sequences b, h, and j contain both of the left-monomer-specific bases diagnostic of the youngest subfamily, PV (also known as HS) (2), while clones 4, 6, 7, 9, and c correspond to precise and clones 2, 3, 13, d, e, o, and n contain a dinucleotide (not found in the aforementioned subfamilies; position 65) characteristic of the older major subfamily. Sequence relatedness among the sc*Alu* cDNAs was found to be higher than expected. For example, four of the seven major representatives are identical except for one base at the extreme 5′ end (clones 13, 2, o, and n). In contrast, the likelihood of finding this degree of identity among independent major sequences deposited in the data base is extremely low. Thus, these four sc*Alu* cDNAs, which contained different-length 5′ poly(dT) tracts (not shown), may represent multiple transcripts of a single major subfamily gene (see discussion).

sc*Alu*s appear to be enriched in young *Alu* sequences in so far as PV sc*Alu* cDNAs comprised ~30% of the population (see Discussion) even though PV-specific sequences reportedly account for only 0.2% of human *Alu* repeats (30). Precise-specific *Alu* repeats comprise ~10% of human DNA but 33% of sc*Alu* cDNAs (Fig. 2). We conclude that sc*Alu* RNAs correspond to Pol III transcripts of the left monomeric components of three *Alu* SINE subfamilies and that young *Alu* sequences are overrepresented in sc*Alu* RNA (see discussion).

Despite the sequence divergence among sc*Alu* RNA sequences, these sequences appear to be enriched in the *Alu* domain RNA secondary structure (46) as predicted by the Cold Spring Harbor minimal-free-energy algorithm used by the Zuker program (57), since ~80% have the potential to form this motif, compared with ~20% for *Alu* sequences in the data base (data not shown).

In summary, sc*Alu* sequences share characteristics which distinguish them from random *Alu* repeats, indicating that they represent a select subset of *Alu*-homologous RNA. sc*Alu* RNAs contain a 5′ end coincident with Pol III initiation, are enriched in a specific secondary structure, and are efficiently compartmentalized to the cytoplasm (7, 30). Although one can propose several postulates to account for these data, a simple interpretation is that sc*Alu* RNAs are specific products of Pol III transcription and are not derived from random sequences found interspersed in high-molecular-weight RNA. sc*Alu* RNAs accumulate to approximately 10^3 to 10^4 copies per (HeLa) cell, a level comparable to the low-abundance small nuclear U-rich RNAs (not shown).

Murine × human somatic cell hybrids express human *Alu* transcripts. Although *Alu* and B1 repeats represent species-specific elements, several observations suggested that their expression utilizes evolutionarily conserved mechanisms: (i) sc*Alu* and scB1 RNAs are expressed in a variety of tissue

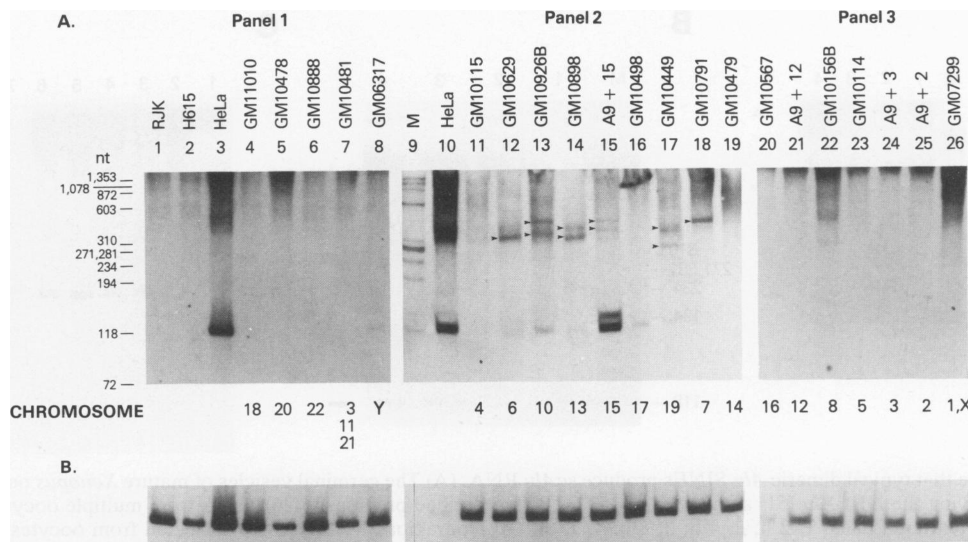

FIG. 3. Alu blot analysis of rodent × human hybrid cell RNA. (A) Total RNA from hybrid cells was fractionated by denaturing 5.5% PAGE, transferred to nylon, and probed with a ^{32}P-end-labeled oligonucleotide DNA complementary to bases 65 to 88 of *Alu* (7). Identities of cell lines are indicated above the lanes, and the human chromosome content for each cell line is indicated below the lanes. RJK and H615 are Chinese hamster and mouse cell lines, respectively; MW refers to *Hae*III-digested, ^{32}P-labeled, ϕX174-denatured size markers indicated in nucleotides on the left. sc*Alu* RNA migrates with the 118-nt DNA size marker. Arrowheads in panel 2 indicate larger *Alu*-homologous RNAs which correspond to *Alu* SINE primary transcripts (see text). (B) The blots shown in panel A were rehybridized with a probe complementary to U1 RNA. The HeLa sample in panel 2 represents cytoplasmic RNA; as expected, it contains very little U1 RNA, a nuclear species.

culture cells (7, 26, 30), (ii) a mouse B1 transfected into various human tissue culture cells was faithfully transcribed into RNA whose termini mapped to the known start and stop sites for Pol III (42, 55), (iii) a mouse B1 expressed in *Xenopus laevis* oocytes was transcribed by Pol III, interacted with conserved proteins, and was faithfully processed to yield an scB1 RNA identical to that found in mouse cells (1, 26, 27), and (iv) mouse and human cells each contain a protein that can bind interchangeably with either scB1 or sc*Alu* RNAs in vitro (7).

To explore the hypothesis that sc*Alu* RNAs are derived from various dispersed loci as suggested by the multiple sc*Alu* cDNA sequences (Fig. 2), we took advantage of a panel of murine × human somatic cell hybrids that retain single human chromosomes by subjecting them to sensitive high-resolution *Alu* RNA blot analysis (7, 26). Although scB1 and sc*Alu* RNAs share sequence homology, they are readily distinguishable by size and probe specificity (7). We used an oligonucleotide DNA probe that is complementary to the *Alu* left monomer (7).

First, we demonstrated the human-specific nature of sc*Alu* RNA by comparing Chinese hamster ovary, mouse, and HeLa cells. As expected, only the human cells contained the sc*Alu* RNA species (Fig. 3A, lanes 1 to 3). This distinct species is routinely detected in HeLa and other human-derived cells studied in our laboratory (data not shown). Although we discovered that several hybrids expressed sc*Alu* RNA, most of these did so at low but detectable levels compared with HeLa cells, suggesting that they each expressed a fraction of the total sc*Alu* RNAs in HeLa cells. In contrast, two different chromosome 15-retaining hybrids reproducibly yielded relatively high levels (lane 15 and data not shown). Also as expected, cytoplasmic RNA preparations from HeLa cells (7) (lane 10) and chromosome 15-retaining cells (not shown) also exhibited high levels of sc*Alu* RNA, while their nuclei contained little or none but revealed the expected small nuclear RNAs (data not shown). The RNA blots shown in Fig. 3A were then hybridized to a probe complementary to U1 RNA (Fig. 3B) as a control for variations in sample loading. Although the hybrid cell lines were previously characterized, we verified their human chromosomal content by Giemsa-11 staining (3) (see Materials and Methods).

To the extent that chromosomes 15 and 17 are relatively small but encoded more sc*Alu* RNA than did larger chromosomes (e.g., chromosomes 1 and X), there was no apparent correlation between the amount of human DNA retained by the hybrids and the amount of sc*Alu* RNA expressed. This finding suggests that some *Alu* DNA loci are more actively transcribed than others and that the total number of expressed loci is small. Although the exact mechanism of biogenesis is unknown, this analysis demonstrated that multiple dispersed loci encode sc*Alu* RNA in the hybrid cells. These loci reside on several human chromosomes and include chromosomes 6, 10, 13, 15, 17, 19, and Y. Although not visualized in Fig. 3, longer exposures of these and other blots as well as phosphor-storage imaging also revealed the sc*Alu* RNA in hybrids that retain human chromosomes 4, 7, 8, 9, and 12. For the remaining 12 negative chromosomes, we cannot rule out the possibility that low levels were undetectable by our assay.

Although internal promoters for Pol III initiation are contained within *Alu* SINEs, the $(dT)_4$ Pol III transcriptional termination signal is not but may be found in 3′ flanking DNA. As a result, transcripts of individual *Alu* SINEs terminate at variable distances from initiation, often within 500 nt but sometimes more distantly (10, 12, 25, 43). A diffuse group of bands of 300 to 500 nt, the expected size range of primary transcripts from *Alu* SINEs, can be reproducibly detected on HeLa cell *Alu* RNA blots (7) (Fig. 3, lane 10). This finding suggests that multiple primary *Alu* SINE transcripts of variable length exist in HeLa cells

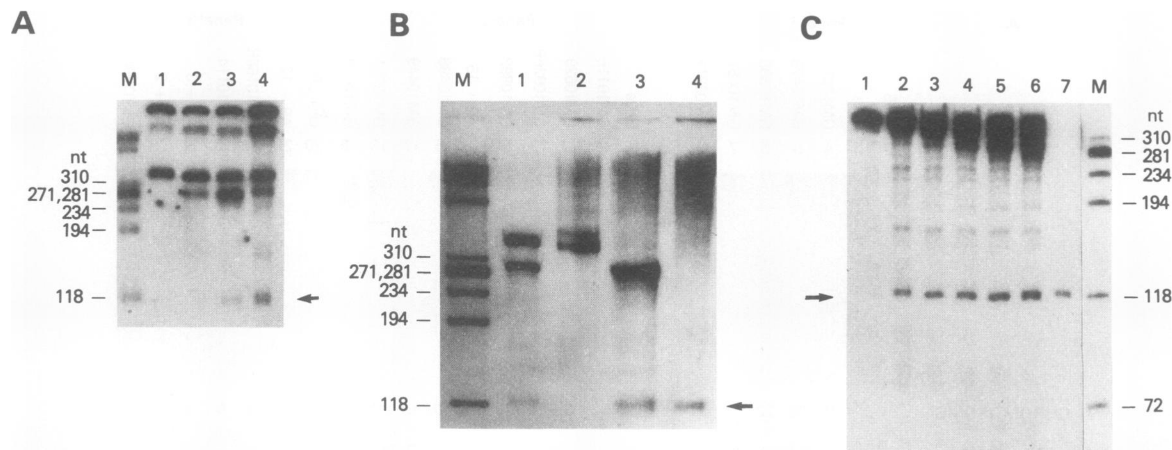

FIG. 4. Evidence that typical dimeric *Alu* SINEs produce sc*Alu* RNA. (A) The germinal vesicles of mature *Xenopus* oocytes were injected with a plasmid carrying the NF1 *Alu* (51) along with [α-^{32}P]GTP as described previously (26). RNA from multiple oocytes was prepared at each time point thereafter. Lanes: 1, 2 h; 2, 4 h; 3, 8 h; 4, 15 h. (B) Nonradioactive RNA was collected from oocytes 15 h postinjection, fractionated by oligo(dT) chromatography, separated by PAGE, transferred to nylon, and probed as in Fig. 3A. Lanes: 1, total RNA; 2, poly(A)$^+$ RNA; 3, poly(A)$^-$ RNA; 4, HeLa cell RNA. Lanes M, size markers as in Fig. 3A. (C) *Alu* RNA synthesized in vitro which corresponds to a Pol III-initiated and terminated transcript was subsequently labeled on its 5′ end and subjected to HeLa cell nuclear extract in reactions containing 4 Units of RNasin per μl as previously described (26). At various times thereafter, equal amounts were removed, and RNA was purified and analyzed by PAGE and autoradiography. Lanes: 1, no extract, no incubation; 2, 10 min; 3, 30 min; 4, 60 min; 5, 90 min; 6, 120 min; 7, 115-nt synthetic RNA made from an *Alu* left monomer (7) and used as a size marker; M; DNA size markers as in Fig. 3A. Arrows indicate positions of sc*Alu* RNA.

possibly as precursors of the various sc*Alu* RNA species. Hybrid cell RNA samples that produced prominent bands in this region were analyzed on the same blot for direct size comparison (Fig. 3A, panel 2). Interestingly, most of the hybrid cells that expressed the sc*Alu* RNA species also contained distinct band(s) in the 300- to 500-nt range (arrowheads in Fig. 3A, panel 2). This result indicated that variable-length *Alu* SINE primary transcripts are produced by multiple loci and also suggested a precursor-product relationship between the larger RNAs and the sc*Alu* species.

In summary, multiple *Alu* primary transcripts as well as sc*Alu*s appear to have been resolved into individual species in the hybrid cells. The size, chromosomal distribution, and heterogeneous nature of the 300- to 500-nt Alu RNAs suggest that they are transcribed from SINEs.

Typical dimeric *Alu* SINEs can produce sc*Alu* RNA. Next, we examined whether cloned *Alu* SINEs expressed in microinjected frog oocytes could produce the sc*Alu* RNA species. This approach was useful for studying Pol III-mediated expression of a B1 SINE and in that case was shown to involve accurate posttranscriptional processing to generate scB1 RNA identical to what is found in mouse cells (1, 26–28). For Fig. 4A, we injected a plasmid carrying the recently transposed NF1 human *Alu* (51) along with [α-^{32}P]GTP into oocyte nuclei and isolated RNA at 2, 4, 8, and 15 h thereafter. The sequence of this *Alu* locus predicts a primary transcript initiated and terminated by Pol III to be ~350 nt and to include a >40-nt poly(A) tract (51). This predicted size agrees with the size of the transcript generated in vitro by using HeLa nuclear extract (not shown). By 2 h after injection, a discrete band of ~350 nt was produced (Fig. 4A, lane 1). At later times, two additional bands appeared at 280 and 118 nt (lanes 2 to 4). The kinetic profiles and small sizes of these RNAs suggested their derivation from the primary RNA by posttranscriptional processing. The pattern of primary transcript, 280 nt, and 118-nt RNAs was also seen after microinjection of another cloned *Alu* SINE of the precise subfamily (15a).

A Pol III-synthesized ~350-nt primary transcript of the injected *Alu* should contain a >40-nt poly(A) region, while the 280- and 118-nt species would correspond to *Alu* dimeric and left monomeric RNAs that have lost their poly(A). To corroborate this interpretation, nonradioactive oocyte-synthesized *Alu* RNA was fractionated by oligo(dT) chromatography and subjected to *Alu* RNA blot analysis using the left-monomer-specific probe (Fig. 4B). Unfractionated RNA is shown in lane 1. As expected, the 350-nt primary transcript was retained by oligo(dT) (lane 2), while the 280- and 118-nt *Alu* species were not but could be recovered in the flowthrough fraction (lane 3). For comparison, HeLa RNA was loaded in lane 4. This hybridization demonstrated that the RNAs produced by injected oocytes were from the *Alu* sequence and that a transposed *Alu* SINE can produce a left-monomer-size, poly(A)$^-$ *Alu* RNA that corresponds to HeLa sc*Alu*. In a control hybridization of this blot using a right-monomer-specific probe, a discrete band of ~150 nt (the size of the right *Alu* monomer) became detectable but only at 5 to 10% of the level of left-monomer sc*Alu* RNA relative to the 350- and 280-nt bands which were detected by either probe. No discrete bands were detectable in the HeLa sample with use of the right-monomer probe, however, even upon prolonged exposure (data not shown). The data indicated that sc*Alu* RNA derived from the left monomer is preferentially accumulated in vivo in microinjected frog oocytes and in HeLa cells.

Finally, to directly demonstrate that a left-monomer *Alu* RNA which corresponds to the sc*Alu* species characterized here can be generated by 3′ processing, we monitored the fate of a purified in vitro-synthesized *Alu* primary transcript after incubation with HeLa cell nuclear extract. This method was useful for studying 3′-terminus formation of scB1 RNA (26, 28). The substrate RNA used for this analysis corresponds to a Pol III-initiated transcript derived from the human alpha-fetoprotein locus (38). After synthesis, the RNA was dephosphorylated and 5′ end labeled by polynucleotide kinase and [γ-^{32}P]ATP. Figure 4C shows that the

sc*Alu* RNA was generated from the larger transcript as a discrete species by 3′ processing. A synthetic *Alu* left-monomer RNA of 115 nt was used as a size marker in lane 7. In separate experiments, this and other primary transcript *Alu* RNAs were uniformly labeled and subjected to nuclear extract; they also yielded the sc*Alu*-size band as the major discrete accumulated species (26a). The discrete size and accumulation of the small RNA over time reflect accurate 3′ processing and stability, respectively. A partial characterization of this HeLa extract-mediated *Alu* RNA processing activity revealed that (i) it is dependent on divalent cations, (ii) the rate of sc*Alu* RNA accumulation is substrate dependent, and (iii) processing occurs in the presence of large amounts of RNase inhibitors (data not shown). Since processing requires Mg^{2+} in these reactions and occurs in the presence of RNasin, it is unlikely that processing is a result of contamination with RNase A. Although the mechanism, nature, and specificity of the processing activity are unknown at present, these results nevertheless demonstrated that *Alu* SINE transcripts are capable of sc*Alu* RNA production by 3′ processing and that HeLa and other cells contain the components necessary to generate a stable left-monomer sc*Alu* RNA.

DISCUSSION

We conclude that sc*Alu* RNA-encoding sequences reside at multiple dispersed loci. Our results indicate that sc*Alu* RNAs are synthesized by Pol III from a subset of transcriptionally active *Alu* SINEs of various subfamily identities. Their relatively high degree of sequence homology with the left monomeric components of contemporary dimeric *Alu*s makes it unlikely that sc*Alu*s are derived from free-standing left monomers (19, 34, 36, 37). Although the data do not exclude the possibility that transcriptional termination within the A-rich spacer (14) accounts for some of the sc*Alu*s, the kinetics observed with the microinjected NF1 *Alu* (Fig. 4A) argue that this is unlikely. Although the cumulative data indicate that sc*Alu* RNAs arise by posttranscriptional processing of *Alu* SINE primary transcripts, this remains to be demonstrated in mammalian cells. Thus, although the mechanism of biogenesis of sc*Alu* RNA is presently unknown, it is clear that human tissue culture cells specifically accumulate this species of cytoplasmic RNA. Our analyses indicates that those *Alu* sequences which have the sequence potential to form the *Alu* domain secondary structure are preferentially expressed as small cytoplasmic RNAs. It is noteworthy that the first four nucleotides of the A-box promoter participate in base pairing within the first stem-loop motif of the *Alu* domain cruciform structure. Therefore, it is difficult to know at this point whether selection is at a transcriptional or posttranscriptional level. Whether posttranscriptional mechanisms such as nucleocytoplasmic transport, RNA stability, or the ability to form the *Alu* domain secondary structure contribute to the selective accumulation of the sc*Alu* subset is also presently unknown.

The overrepresentation of PV-specific sc*Alu* cDNAs was confirmed by two approaches: (i) sc*Alu* cDNA clones were subjected to restriction analysis with *Msp*I, an enzyme that cuts PV-specific but not other *Alu*s at position 89 (40); 33% of clones yielded the PV-specific restriction pattern. Three of these were sequenced (Fig. 2) and found to contain the PV-specific A residue at position 96, confirming their subfamily identity. Also, HeLa cell sc*Alu* RNA was calibrated by Northern blot analysis to PV and precise sequence RNAs synthesized in vitro from cloned *Alu*s. PV-specific and nondiscriminatory probes were hybridized independently. Subsequent analysis indicated that ~30% of sc*Alu* RNA in HeLa cells was PV specific. These analyses suggest that the relative frequency with which sc*Alu* cDNAs were obtained reflects the approximate relative levels of the corresponding sc*Alu* RNA in vivo. For example, although clones 13, 2, o, and n share identical *Alu* sequences, the lengths of their 5′ poly(dT) tracts were significantly different from each other (data not shown), consistent with the notion that they represent independently derived cDNAs, not just multiplicate representatives of a single-transcript molecule.

A poly(A)$^-$ 280-nt *Alu* RNA appears as a prominent band in injected oocytes with kinetics suggesting that it may be a processing intermediate (Fig. 4A and B). However, this species accumulates to a lesser degree in HeLa and other human-derived cells and can also be visualized by using a PV-specific probe (not shown). It may be of interest to note that a band of ~280 nt is also detectable in chromosome 19-retaining hybrid cells (Fig. 3A).

Although the *Alu* insertions that contributed variability to the human genome were inherited through the germ line, it is plausible that transposition is not limited to this tissue and that *Alu* insertion may also occur in somatic cells (9, 24). Identification of RNAs that correspond to primary transcripts of *Alu* sequences (7, 30) (Fig. 3) suggests the possibility that one or more of the sequences from which they were transcribed hold some potential for transposition (30). However, posttranscriptional determinants such as sequence specificity, relative stability, RNA folding, and others may affect retroposition (39, 47). Whether or not *Alu* source gene transcripts, once identified, would be susceptible to conversion to sc*Alu* RNA awaits identification and cloning of an active source gene.

Results of others indicate that multiple *Alu* sequences may serve as sources of new transpositions (17, 23, 29). According to a current model of retrotransposition, an *Alu* source gene may be an inserted element which acquired flanking sequences that promote transposition (39). Our data show that multiple *Alu* loci are transcriptionally active in tissue culture cells (Fig. 2 and 3) and are consistent with such a model. Although we wish to emphasize that the functional relationship, if any, between the *Alu* RNAs identified here and *Alu* source sequences is completely unknown. Yet it is possible that one or more of the *Alu* transcripts identified here is also active in the germ line and fulfills other requirements for transposition. In an attempt to map candidate loci from which the *Alu* gene(s) mediate contemporary human transpositions, hybrid cell RNA blots are being probed with PV- and NF1-specific oligonucleotides. Once identified, probes to the associated 3′ unique sequence can be used on germ line RNA.

The sequences of sc*Alu* RNAs indicate that many are derived from elements that have undergone sequence drift away from their subfamily consensus while maintaining both transcriptional competence and potential to fold into the *Alu* secondary structure. The fact that sc*Alu* and scB1 RNAs accumulate to significant levels suggests that they do not exist as naked molecules within cells but that they become associated with proteins or other factors. The ability to identify a protein that binds sc*Alu* and scB1 RNA illustrates this potential (7). The existence of a family of multiple sc*Alu* RNA species of sequence microheterogeneity demonstrates the capacity and versatility of dispersed elements to provide the cell with a variety of RNAs that may have the potential to interact with cellular factors. The preservation by some of the sc*Alu* species of a sequence and secondary structure

motif also found in the translational control domain of signal recognition particle RNAs from several organisms (22, 49) further suggests that such dispersed elements may come to interact with preexisting specific biochemical pathways, perhaps to interfere with or modify existing functions, or to take on new activities. This contemporary perspective, which accounts for dispersed elements not as useless junk but as a potential source of new sequences from which some may be recruited into novel functions during evolution, is shared by several biologists (5, 6, 15, 56). It may be of interest to note that although most sc*Alu* RNAs maintain the *Alu* cruciform structure, some fold into variable structures at the other end of the molecule, creating the potential for linking a translation control domain with a limited number of other RNA structures as was suggested for scB1 RNA (26).

The unexpectedly high level of chromosome 15-encoded sc*Alu* RNA may be due to a relatively high level of transcription from a single locus, for example as discussed earlier with regard to the overrepresented major subfamily sc*Alu* cDNA sequence (Fig. 2). This interpretation is consistent with the fact that a PV-specific oligonucleotide probe detects much less sc*Alu* RNA from chromosome 15 cells than does the nondiscriminatory probe in comparison to the other samples on the same blot. However, the data are also consistent with another explanation. Examination of scB1 RNA in the rodent × human cells also reveals chromosome 15-specific overexpression, suggesting the important possibility that chromosome 15 encodes a *trans*-acting factor whose expression leads to elevated levels of these scRNAs (7a). Because the ratio of small cytoplasmic to the larger *Alu* transcript was highest in the chromosome 15 hybrid, it is tempting to speculate that this putative factor is involved in *Alu* transcript metabolism and that defects in such a factor, if it is involved in *Alu* source gene transcript stability and expressed in the germ line, may lead to an increased frequency of heritable *Alu* transpositions.

ACKNOWLEDGMENTS

We thank M. Wallace and F. Collins for the NF1 *Alu*, O. Pereira-Smith for hybrid cell lines, B. Howard, C. Schmid, M. Singer, and an anonymous reviewer for helpful comments, and M. L. Lanigan for preparation of the manuscript. We also thank J. Brosius for sharing information prior to publication.

K.H. was supported by an Intramural Research Training Award from the NICHHD.

REFERENCES

1. **Adeniyi-Jones, S., and M. Zasloff.** 1985. Transcription, processing and nuclear transport of a B1 *Alu* RNA species complementary to an intron of the murine α-fetoprotein gene. Nature (London) **317:**81–84.
2. **Batzer, M. A., G. E. Kilroy, P. E. Richard, T. H. Shaikh, T. D. Desselle, C. L. Hoppens, and P. L. Deininger.** 1990. Structure and variability of recently inserted Alu family members. Nucleic Acids Res. **18:**6793–6798.
3. **Bobrow, M., and J. Cross.** 1974. Differential staining of human and mouse chromosomes in interspecific cell hybrids. Nature (London) **251:**77–79.
4. **Britten, R. J., W. F. Baron, D. B. Stout, and E. H. Davidson.** 1988. Sources and evolution of human *Alu* repeated sequences. Proc. Natl. Acad. Sci. USA **85:**4770–4774.
5. **Brosius, J.** 1991. Retroposons—seeds of evolution. Science **251:**753.
6. **Brosius, J., and S. J. Gould.** 1992. On "genomenclature": a comprehensive (and respectful) taxonomy for pseudogenes and other "junk DNA." Proc. Natl. Acad. Sci. USA **89:**10706–10710.
7. **Chang, D.-Y., and R. J. Maraia.** 1993. A cellular protein binds B1 and Alu small cytoplasmic RNAs *in vitro*. J. Biol. Chem. **268:**6423–6428.

7a. **Chang, D.-Y., et al.** Unpublished data.

8. **Daniels, G. R., and P. L. Deininger.** 1991. Characterization of a third major SINE family of repetitive sequences in the galago genome. Nucl. Acids Res. **19:**1649–1656.

8a. **Driscoll, C., R. Maraia, and B. Howard.** Unpublished data.

9. **Economou-Pachnis, A., and P. N. Tsichlis.** 1985. Insertion of an Alu SINE in the human homologue of the *Mlvi*-2 locus. Nucleic Acids Res. **13:**8379–8387.
10. **Elder, J. T., J. Pan, C. H. Duncan, and S. M. Weissman.** 1981. Transcriptional analysis of interspersed repetitive polymerase III transcription units in human DNA. Nucleic Acids Res. **9:**1171–1189.
11. **Frohman, M. A., M. K. Dush, and G. R. Martin.** 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85:**8998–9002.
12. **Fuhrman, S. A., P. L. Beininger, P. LaPorte, T. Friedmann, and E. P. Geiduschek.** 1981. Analysis of transcription of the human Alu family ubiquitous repeating element by eukaryotic RNA polymerase III. Nucleic Acids Res. **9:**6439–6456.
13. **Gross, D. S., K. W. Collins, E. M. Hernandez, and W. T. Garrard.** 1988. Vacuum blotting: a simple method for transferring DNA from sequencing gels to nylon membranes. Gene **74:**347–356.
14. **Hess, J., C. Perez-Stable, G. J. Wu, B. Weir, I. Tinoco, Jr., and C.-K. J. Shen.** 1985. End-to-end transcription of an *Alu* family repeat: a new type of polymerase-III-dependent terminator and its evolutionary implication. J. Mol. Biol. **184:**7–21.
15. **Howard, B. H., and K. Sakamoto.** 1990. Alu interspersed repeats: selfish DNA or a functional gene family. New Biol. **2:**759–770.

15a. **Hsu, K., and R. Maraia.** Unpublished data.

16. **Jelinek, W. R., and C. W. Schmid.** 1982. Repetitive sequences in eukaryotic DNA and their expression. Annu. Rev. Biochem. **51:**813–844.
17. **Jurka, J., and A. Milosavljevic.** 1991. Reconstruction and analysis of human Alu genes. J. Mol. Evol. **32:**105–121.
18. **Jurka, J., and T. Smith.** 1988. A fundamental division in the *Alu* family of repeated sequences. Proc. Natl. Acad. Sci. USA **85:**4775–4778.
19. **Jurka, J., and E. Zuckerkandl.** 1991. Free left arms as precursor molecules in the evolution of Alu sequences. J. Mol. Evol. **33:**49–56.
20. **Labuda, D., D. Sinnett, C. Richer, J.-M. Deragon, and G. Striker.** 1991. Evolution of mouse B1 repeats: 7SL RNA folding pattern conserved. J. Mol. Evol. **32:**405–414.
21. **Labuda, D., and G. Striker.** 1989. Sequence conservation in Alu evolution. Nucleic Acids Res. **17:**2477–2491.
22. **Larsen, N., and C. Zwieb.** 1991. SRP-RNA sequence alignment and secondary structure. Nucleic Acids Res. **19:**209–215.
23. **Leeflang, E. P., V. M. Liu, C. Hashimoto, P. V. Choudary, and C. W. Schmid.** 1992. Phylogenetic evidence for multiple Alu source genes. J. Mol. Evol. **35:**7–16.
24. **Lin, C. S., D. A. Goldthwait, and D. Samols.** 1988. Identification of Alu transposition in human lung carcinoma cells. Cell **54:**153–159.
25. **Manley, J. L., and M. T. Colozzo.** 1982. Synthesis *in vitro* of an exceptionally long RNA transcript promoted by an *Alu*I sequence. Nature (London) **300:**376–379.
26. **Maraia, R.** 1991. The subset of mouse B1 (Alu-equivalent) sequences expressed as small processed cytoplasmic transcripts. Nucleic Acids Res. **19:**5695–5702.

26a. **Maraia, R.** Unpublished data.

27. **Maraia, R., M. Zasloff, P. Plotz, and S. Adeniyi-Jones.** 1988. Pathway of B1-*Alu* expression in microinjected oocytes: *Xenopus laevis* proteins associated with nuclear precursor and processed cytoplasmic RNAs. Mol. Cell. Biol. **8:**4433–4440.
28. **Maraia, R. J., D.-Y. Chang, A. P. Wolffe, R. L. Vorce, and K. Hsu.** 1992. The RNA polymerase III terminator used by a B1-*Alu* element can modulate 3′ processing of the intermediate RNA product. Mol. Cell. Biol. **12:**1500–1506.

29. **Matera, A. G., U. Hellmann, M. F. Hintz, and C. W. Schmid.** 1990. Recently transposed Alu repeats result from multiple source genes. Nucleic Acids Res. **18:**6019–6023.
30. **Matera, A. G., U. Hellmann, and C. W. Schmid.** 1990. A transpositionally and transcriptionally competent Alu subfamily. Mol. Cell. Biol. **10:**5424–5432.
31. **Mitchell, G. A., D. Labuda, G. Fontaine, J. M. Saudubray, J. P. Bonnefont, S. Lyonnet, L. C. Brody, G. Steel, C. Obie, and D. Valle.** 1991. Splice-mediated insertion of an *Alu* sequence inactivates ornithine δ-aminotransferase: a role for *Alu* elements in human mutation. Proc. Natl. Acad. Sci. USA **88:**815–819.
32. **Murphy, M. H., and F. E. Baralle.** 1983. Directed semisynthetic point mutational analysis of an RNA polymerase III promoter. Nucleic Acids Res. **11:**7695–7700.

32a. **NIGMS Human Genetic Mutant Cell Repository.** 1992. 1992/1993 catalog of cell lines. NIH publication 92-2011. National Institute of General Medical Sciences, Bethesda, Md.

33. **Ning, Y., M. Lovell, L. Taylor, and O. M. Pereira-Smith.** 1992. Isolation of monochromosomal hybrids following fusion of human diploid fibroblast-derived microcells with mouse A9 cells. Cytogenet. Cell Genet. **60:**79–80.
34. **Quentin, Y.** 1988. The Alu family developed through successive waves of fixation closely connected with primate lineage history. J. Mol. Evol. **27:**194–202.
35. **Quentin, Y.** 1989. Successive waves of fixation of B1 variants in rodent lineage history. J. Mol. Evol. **28:**299–305.
36. **Quentin, Y.** 1992. Fusion of a free left Alu monomer and a free right Alu monomer at the origin of the Alu family in the primate genomes. Nucleic Acids Res. **20:**487–493.
37. **Quentin, Y.** 1992. Origin of the Alu family: a family of Alu-like monomers gave birth to the left and the right arms of the Alu elements. Nucleic Acids Res. **20:**3397–3401.
38. **Ryan, S. C., and A. Dugaiczyk.** 1989. Newly arisen DNA repeats in primate phylogeny. Proc. Natl. Acad. Sci. USA **86:**9360–9364.
39. **Schmid, C., and R. Maraia.** 1992. Transcriptional regulation and transpositional selection of active SINE sequences. Curr. Opin. Genet. Dev. **2:**874–882.
40. **Schmid, C. W.** 1991. Human Alu subfamilies and their methylation revealed by blot hybridization. Nucleic Acids Res. **19:** 5613–5617.
41. **Schmid, C. W., and C. K. J. Shen.** 1985. The evolution of interspersed repetitive DNA sequences in mammals and other vertebrates, p. 323–358. *In* R. J. MacIntyre (ed.), Molecular evolutionary genetics. Plenum, New York.
42. **Scott, R. W., and S. M. Tilghman.** 1983. Transient expression of a mouse α-fetoprotein minigene: deletion analyses of promoter function. Mol. Cell. Biol. **3:**1295–1309.
43. **Shen, C.-K. J., and T. Maniatis.** 1982. The organization, structure, and *in vitro* transcription of Alu family RNA polymerase III transcription units in the human α-like globin gene cluster: precipitation of *in vitro* transcripts by lupus anti-La antibodies. J. Mol. Appl. Genet. **1:**343–360.
44. **Shen, M. R., M. A. Batzer, and P. L. Deininger.** 1991. Evolution of the master Alu gene(s). J. Mol. Evol. **33:**311–320.
45. **Singer, M. F.** 1982. SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. Cell **28:**433–434.
46. **Sinnett, D., C. Richer, J.-M. Deragon, and D. Labuda.** 1991. *Alu* RNA secondary structure consists of two independent 7 SL RNA-like folding units. J. Biol. Chem. **266:**8675–8678.
47. **Sinnett, D., C. Richer, J.-M. Deragon, and D. Labuda.** 1992. Alu RNA transcripts in human embryonal carcinoma cells model of post-transcriptional selection of master sequences. J. Mol. Biol. **226:**689–706.
48. **Slagel, V., E. Flemington, V. Traina-Dorge, H. Bradshaw, Jr., and P. L. Deininger.** 1987. Clustering and sub-family relationships of the Alu family in the human genome. Mol. Biol. Evol. **4:**19–29.
49. **Strub, K., J. B. Moss, and P. Walter.** 1991. Binding sites of the 9- and 14-kilodalton heterodimeric protein subunit of the signal recognition particle (SRP) are contained exclusively in the *Alu* domain of SRP RNA and contain a sequence motif that is conserved in evolution. Mol. Cell. Biol. **11:**3949–3959.
50. **Tiedge, H., W. Chen, and J. Brosius.** Primary structure, neural-specific expression, and dendritic location of human BC200 RNA. J. Neurosci., in press.
51. **Wallace, M. R., L. B. Andersen, A. M. Saulino, P. E. Gregory, T. W. Glover, and F. S. Collins.** 1991. A *de novo Alu* insertion results in neurofibromatosis type 1. Nature (London) **353:**864–866.
52. **Watson, J. B., and J. G. Sutcliffe.** 1987. Primate brain-specific cytoplasmic transcript of the Alu repeat family. Mol. Cell. Biol. **7:**3324–3327.
53. **Weiner, A. M., P. L. Deininger, and A. E. Estratiadis.** 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu. Rev. Biochem. **55:**631–662.
54. **Willard, C., H. T. Nguyen, and C. W. Schmid.** 1987. Existence of at least three distinct Alu subfamilies. J. Mol. Evol. **26:**180–186.
55. **Young, P. R., R. W. Scott, D. H. Hamer, and S. M. Tilghman.** 1982. Construction and expression in vivo of an internally deleted mouse α-fetoprotein gene: presence of a transcribed Alu-like repeat within the first intervening sequence. Nucleic Acids Res. **10:**3099–3116.
56. **Zuckerkandl, E., G. Latter, and J. Jurka.** 1989. Maintenance of function without selection: *Alu* sequences as "cheap genes." J. Mol. Evol. **29:**504–512.
57. **Zuker, M.** 1989. Computer prediction of RNA structure. Methods Enzymol. **180:**262–288.