MOLECULAR AND CELLULAR BIOLOGY, Sept. 1987, p. 3324–3327
0270-7306/87/093324-04$02.00/0

Vol. 7, No. 9

Primate Brain-Specific Cytoplasmic Transcript of the Alu Repeat Family†

JOSEPH B. WATSON AND J. GREGOR SUTCLIFFE*

Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037

Received 4 May 1987/Accepted 16 June 1987

A 200-nucleotide RNA homologous to the left monomer of Alu elements was expressed in monkey and human brain and in cell lines but not in nonneural monkey tissues. Similar brain-specific transcription of identifier sequences was observed in rats. Thus, expression of selected repetitive DNA families is a conserved process in mammalian brain.

Members of the Alu repeat family are dispersed throughout the genomes of primates at frequencies ranging from 3×10^5 to 9×10^5 copies per genome (12, 13). On the basis of the presence of apparent RNA polymerase III (Pol III) promoter elements and 3′ A tracts within the 300-base-pair Alu sequences, as well as short direct repeats flanking Alu family members, it was proposed that amplification and dispersal of Alu sequences occurred by retroposition of an ancestral Alu RNA intermediate (14, 19, 30). Alu elements are composed of two modules, the so-called left and right monomers, each of which appears to have evolved from an internally deleted 7SL RNA gene (26). Here we report the detection of a 200-nucleotide monkey brain-specific cytoplasmic poly(A)$^+$ RNA which hybridized at low stringency to the rat identifier (ID) element (22). Sequence analysis showed that the monkey brain-specific RNA is highly homologous to the human Alu left monomer (5, 21).

Small rat brain-specific RNAs BC1 (160 nucleotides) and BC2 (110 nucleotides) have been described previously and were shown to be transcribed from repetitive ID sequences (22, 23). We tested whether ID-homologous transcripts could be detected in monkey brain. Cytoplasmic poly(A)$^+$ RNAs prepared from brain and liver of cynomolgus monkeys and rats were resolved by electrophoresis on a formaldehyde-containing 1.5% agarose gel, transferred to nitrocellulose (25), and hybridized to ^{32}P-labeled ID cDNA probe p2A120 (22). The blot was hybridized in 56% formamide at 37°C as described previously (7) and washed in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.5% sodium dodecyl sulfate at 42°C (low stringency). The ID probe hybridized with an abundant 200-nucleotide RNA present in monkey brain but absent from monkey liver (Fig. 1). This RNA, which we will call monkey BC200 RNA, is slightly larger than the rat brain-specific ID transcripts BC1 and BC2. Monkey BC200 RNA was not detected after higher-stringency washes (equivalent to about 85% homology; data not shown), indicating that it had only partial similarity to the rat ID sequence.

Our strategy for determining the sequence of monkey BC200 RNA was based on its presence in the poly(A)$^+$ RNA fraction. We first determined which of 11 different combinations of the synthetic deoxyoligonucleotide [5′-^{32}P]dT_{12}VX (V, AGC; X, AGCT) would serve as efficient primers for reverse transcriptase extension reactions with monkey brain poly(A)$^+$ RNA as a template (Fig. 2A). Primer extension reactions were performed as described previously (9). Extended cDNA products were resolved by electrophoresis on 6% polyacrylamide–7 M urea sequencing gels. Each primer (1 pmol) was extended to make the long products that run near the top of the gel and presumably represent cDNA copies of mRNAs. The primer dT_{12}CC and, to lesser degrees, primers dT_{12}CT and dT_{12}AC gave rise to abundant cDNA products about 135 nucleotides long. These comigrated with the larger of two cDNAs extended with the primer dT_{12}GG on rat brain RNA. The two rat products (130 to 140 and 100 nucleotides) have been shown to correspond to copies of the 5′ portions of BC1 and BC2 RNAs, respectively (M. A. Brow and J. G. Sutcliffe, manuscript in preparation). Because only one length of abundant extended products was observed in the monkey lanes, we concluded

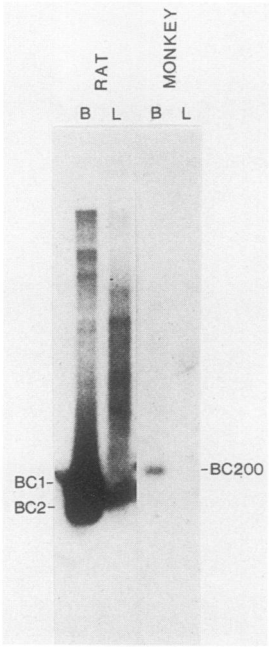

FIG. 1. Northern blot detection of ID hybridization with monkey brain poly(A)$^+$ RNA. A Northern blot of cytoplasmic poly(A)$^+$ RNAs (2, 20) (2 μg) prepared from brain (B) and liver (L) of rats and cynomolgus monkeys was hybridized to ^{32}P-labeled ID cDNA probe p2A120 (22). The positions of BC1 (160 nucleotides) and BC2 (110 nucleotides) RNAs, along with that of monkey BC200 RNA, are marked.

* Corresponding author.

† Publication 4690-MB from the Research Institute of Scripps Clinic.

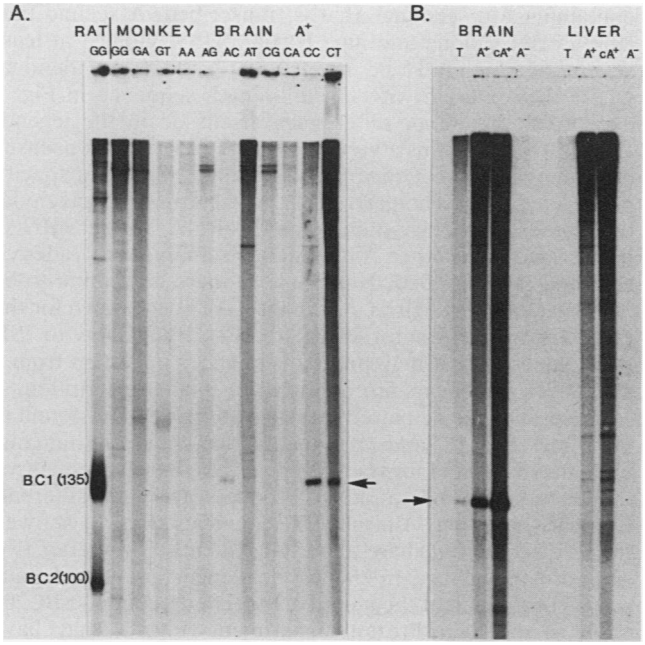

FIG. 2. Primer extension analysis of monkey brain poly(A)$^+$ RNA. (A) Poly(A)$^+$ RNA from a cynomolgus monkey brain total RNA preparation (1, 6) was used as a template for a primer extension reaction with reverse transcriptase and one of 11 different synthetic deoxyoligonucleotide primers of the form [5'-^{32}P]dT$_{12}$VX (New England BioLabs, Inc.). Monkey brain cDNA products which comigrated with cDNA copies of rat BC1 RNAs (5' 130 to 140 nucleotides, first lane) primed with dT$_{12}$GG are indicated by an arrow. (B) Samples from monkey brain or liver total RNA (T), poly(A)$^+$ RNA (A$^+$), cytoplasmic poly(A)$^+$ RNA (cA$^+$), and poly(A)$^-$ RNA (A$^-$) were templates for primer extension reactions with primer dT$_{12}$CC, referred to as oligo I.

tentatively that these corresponded to partial copies of the BC200 RNA that hybridized at low stringency to the rat ID probe. The observation that at least three primers can be extended to 135 nucleotides suggests that there are different populations of monkey BC200 RNAs, consistent with similar observations for rat BC RNAs (Brow and Sutcliffe, in preparation). Only cDNA copies generated with dT$_{12}$CC, referred to here as oligo I, were studied in further detail.

Oligo I was used as a primer with total, poly(A)$^+$, cytoplasmic poly(A)$^+$, and poly(A)$^-$ RNAs from monkey brain and liver (Fig. 2B). It was an efficient primer for synthesis of the 135-nucleotide cDNA for all brain poly(A)$^+$ RNA samples but not for brain poly(A)$^-$ or any liver RNA samples. The sequence of the oligo I-primed cDNA made from the monkey brain poly(A)$^+$ RNA template was determined by scaling up the primer extension reaction and isolating the 135-nucleotide cDNA (approximately 20,000 cpm) by excision from a polyacrylamide gel and subjecting it to the partial chemical degradation sequencing technique (15) (Fig. 3). This provided an apparently unique sequence for the 5' 138 nucleotides of the 200-nucleotide RNA; the remaining 3' sequence may correspond to a poly(A) tract. The accuracy of the primer extension cDNA sequence was confirmed by using the sequence to design an additional 19mer synthetic deoxyoligonucleotide, oligo II (Fig. 3), which was used in colony hybridization experiments (11) at room temperature to isolate complementary clones from a monkey brain cDNA library (G. H. Travis, C. G. Naus, J. H. Morrison, F. E. Bloom, and J. G. Sutcliffe, Neuropharmacology, in press). The sequence of a full-length cDNA clone (pMB12) was determined and found to be identical to the primer extension sequence with an additional long 3' stretch of adenine residues.

Clone pMB12 hybridized on Northern (RNA) blots (Fig. 4A) at high stringency to a 200-nucleotide monkey brain RNA present in total and poly(A)$^+$, but not poly(A)$^-$, samples. The monkey BC200 RNA had slightly slower mobility in the total RNA fraction because of gel overloading, as shown by the control lanes (Fig. 4B), which also had retarded mobility. The RNA was apparently the same as monkey BC200 RNA (Fig. 1) because when the Northern blot of Fig. 1 was reprobed with pMB12, a 200-nucleotide monkey brain RNA indistinguishable from that previously detected with the ID probe was observed (data not shown). Monkey BC200 RNA was not detectable in RNA from monkey liver or spleen (Fig. 4A) nor was it detected in heart, kidney, colon, or skeletal muscle RNA samples (data not shown). Similar amounts of RNA were loaded in each lane, as shown by probing of the blot (Fig. 4B) with a cDNA clone of a ubiquitous RNA. A human brain poly(A)$^+$ RNA which comigrated with the monkey BC200 RNA hybridized to the pMB12 probe; there was no rat RNA which hybridized at high stringency (Fig. 4A). We also detected analogous poly(A)$^+$ RNAs in both HeLa and human 143 osteosarcoma cells with the pMB12 probe (Fig. 4C). In all Northern blot experiments, faint hybridization to a 300-nucleotide poly(A)$^-$ RNA present in all tissues, especially liver, was detected. Further analysis, discussed below, showed this to be 7SL RNA.

A computer search with the BC200 RNA sequence revealed that it was 77% identical to the human Alu left monomer (5, 21) (Fig. 3). The close homology of the BC200 RNA to the Alu sequence explains hybridization to the 7SL

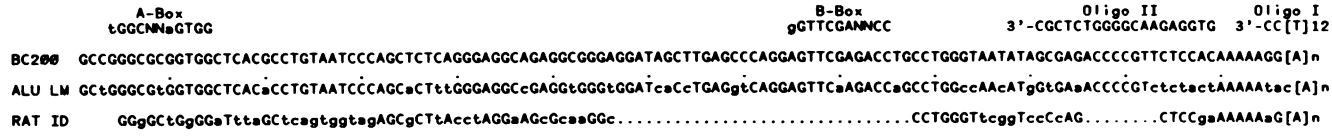

FIG. 3. Comparison of monkey BC200 RNA (shown in cDNA form) with the human Alu left monomer (LM) (5, 21) and the consensus rat ID sequence (24). Consensus RNA Pol III A and B boxes (4, 8) are denoted above the sequence. Nucleotides within the A and B boxes, the Alu left monomer, and the ID sequence homologous to monkey BC200 RNA are denoted by capital letters (A, G, C, and T), and the nonhomologous sequence is denoted by small letters (a, g, c, and t). Gaps in the ID sequence are indicated by dots. Complementary oligo I and oligo II are indicated in the 3' to 5' direction. The sequence of two cDNA clones (pMB12 and pMB4) was identical to the original full-length primer extension sequence (138 nucleotides) plus a 3' stretch of more than 70 adenines. The remaining clones contained 55 to 96 nucleotides of the primer-extended RNA and were identical to it in sequence where they overlapped except that the GG in positions 125 and 126 was AA in seven of the clones. This may be a cloning artifact. The 5' nucleotide could not be directly determined in the primer extension sequence, and a homopolymeric G run added by cloning was found in the cDNA clones. Thus, the single 5' G is assigned based on the known 5' end of 7SL RNA (28, 29).

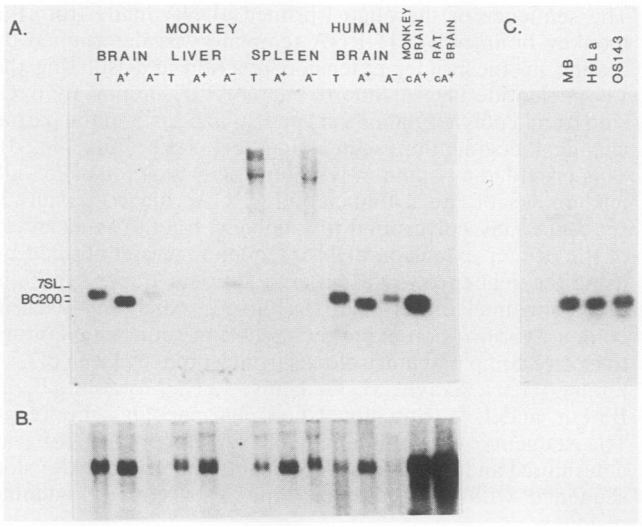

FIG. 4. Northern analysis of tissue distribution of monkey BC200 RNA. (A) A Northern blot of RNA samples from monkey brain, liver, and spleen, as well as human and rat brain, was hybridized to a ^{32}P-labeled 230-base-pair cDNA insert of pMB12. The blot was washed in 0.1× SSC–0.5% sodium dodecyl sulfate at 68°C (high stringency). Approximately 20 μg of total RNA (T) and poly(A)$^-$ RNA (A$^-$) plus 2 μg of either enriched poly(A)$^+$ RNA (A$^+$) or cytoplasmic poly(A)$^+$ RNA (cA$^+$) were loaded in the designated lanes. The positions of monkey BC200 RNA and 7SL RNA (300 nucleotides) are indicated. The slightly slower migration of BC200 RNA in the total RNA fraction was probably due to overloading of the gel. The high-molecular-weight species seen in spleen total and A$^-$ RNA fractions, as well as in all other total and A$^-$ RNA fractions with longer exposure, are most likely heterogeneous nuclear RNA or residual sheared DNA. (B) The blot in panel A was reprobed with a cDNA clone of a ubiquitous RNA (p1B15) (16). (C) Northern blot of poly(A)$^+$ RNA (2 μg) from monkey brain (MB) or cytoplasmic poly(A)$^+$ RNA (2 μg) prepared as described previously (17) from HeLa or human 143 osteosarcoma cells (OS143) (10) was probed with a ^{32}P-labeled cDNA insert of pMB12.

RNA since Alu elements are partially homologous to 7SL and are thought to have evolved from processed 7SL transcripts (27, 31). Thus, the monkey BC200 RNA represents the first detection in vivo of a stable Alu transcript. The existence of an Alu transcript was predicted by the retroposition model (14, 19, 30), but it is unknown to what extent BC200 RNA resembles the ancestral Alu RNA intermediate that retroposed throughout the genome. On the basis of the abundance of the monkey BC200 RNA, the presence in its sequence of RNA Pol III A and B boxes (Fig. 3), and the position of the 5′ end of the RNA relative to the Pol III A box, it seems most likely that this RNA is the product of brain-specific Pol III transcription. Previous work suggests that Pol III transcripts from the human β-globin gene can be polyadenylated (3). However, since Alu sequences are known to be transcribed as components of Pol II heterogeneous nuclear RNA (21), there remains the possibility that posttranscriptional processing of heterogenous nuclear RNA generates BC200 RNAs.

Are all Alu repeats transcribed in primate brain? Three of the dT$_{12}$VX primers we studied gave rise to extension products of 135 nucleotides, suggesting that the monkey BC200 RNA population is heterogeneous and thus is transcribed from at least a few genes. However, we found no evidence of heterogeneity in the dT$_{12}$CC-primed cDNA sequence, suggesting that there is only one dT$_{12}$CC-containing Alu sequence that is transcribed. A second full-length cDNA clone and nine partial cDNA clones, at least several of which must be independent isolates, are identical where they overlap with the full-length sequence in Fig. 3 except for the minor differences discussed in the legend. BC200 RNA contains a variant (AATATA) of the consensus polyadenylation sequence (AATAAA) (18) but does specify polyadenylation. Therefore, it is possible that polyadenylation confers particular stability on this RNA, whereas RNAs transcribed from other Alu elements remain unpolyadenylated and are degraded. Nonetheless, there are bound to be Alu pseudogenes such as those that have been shown for the proposed Alu progenitor 7SL RNA (28). By analogy to 7SL RNA genes (29), it may be that sequences upstream from a subset of Alu repeats are required for their transcription.

The brain-specific pattern of Alu transcription we found in vivo for monkeys (and presumably also occurs in humans) parallels that found for the ID repeat (23), but the sequences of the monkey and rat RNA products are very different. Comparison of the Alu and ID sequences (Fig. 3) showed that the sequences share about 60% identity only after two gaps are introduced in the ID sequence, and the A tails account for much of the identity. Indeed, detection of BC200 RNA in the hybridization experiment (Fig. 1) may have resulted in part from the fortuitous occurrence of a stretch of 12 nucleotides in the region upstream from the ID sequence in cDNA clone p2A120 that matches nucleotides 61 to 72 of BC200 RNA. The primer extension experiment (Fig. 2) indicated that BC200 was the only abundant small poly(A)$^+$ RNA in the monkey brain RNA sample. Thus, the nonsimilarity between the BC200 and ID sequences suggests that neither rodent nor primate BC RNA performs a conserved function. However, their pattern of brain-specific transcription is conserved and therefore may contribute to the normal operation of mammalian brain.

We especially thank Patria Danielson for preparing cytoplasmic RNA samples; Mary Ann D. Brow for advice on primer extension experiments; Gabriel Travis for generously providing human RNA samples and a monkey brain cDNA library; Randy McKinnon for his helpful suggestions in planning experiments and his computer expertise; Peter Staeheli, Frank Burton, Sonja Forss-Petter, and Ute Hochgeschwender for their critical comments; Jürgen Brosius for recognizing the upstream p2A120 sequence relationship; and Linda Elder for preparation of the manuscript.

This work was supported in part by Public Health Service grants GM32355 and NS22347 from the National Institutes of Health. J.B.W. is a fellow of the Huntington's Disease Society of America, Inc.

LITERATURE CITED

1. **Auffray, C., and F. Rougeon.** 1980. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. Eur. J. Biochem. **107**:303–314.
2. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
3. **Carlson, D. P., and J. Ross.** 1983. Human β-globin promoter and coding sequences transcribed by RNA polymerase III. Cell **34**:857–864.
4. **Ciliberto, G., F. Raugei, F. Constanzo, L. Dente, and R. Cortese.** 1983. Common and interchangeable elements in the promoters of genes transcribed by RNA polymerase III. Cell **32**:725–733.
5. **Deininger, P. L., D. J. Jolly, C. M. Rubin, T. Friedmann, and C. W. Schmid.** 1981. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. J. Mol. Biol. **151**:17–33.
6. **Dierks, P., A. VanOoyen, N. Mantei, and C. Weissmann.** 1981.

DNA sequences preceding the rabbit β-globin gene are required for formation in mouse L cells of β-globin RNA with the correct 5′ terminus. Proc. Natl. Acad. Sci. USA **78**:1411–1415.
7. **Forss-Petter, S., P. Danielson, and J. G. Sutcliffe.** 1986. Neuron-specific enolase: complete structure of rat mRNA, multiple transcriptional start sites, and evidence suggesting post-transcriptional control. J. Neurosci. Res. **16**:141–156.
8. **Galli, G., H. Hofstetter, and M. L. Birnstiel.** 1981. Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature (London) **294**:626–631.
9. **Giorgi, C., B. Blumberg, and D. Kolakofsky.** 1983. Sendai virus contains overlapping genes expressed from a single mRNA. Cell **35**:829–836.
10. **Graham, F. L., S. Bacchetti, R. D. McKinnon, C. Stanners, B. Cordell, and H. M. Goodman.** 1980. Transformation of mammalian cells with DNA using the calcium technique, p. 3–25. *In* R. Baserga, C. Croce, and G. Robera (ed.), Introduction of macromolecules into viable mammalian cells. Progress in clinical and biological research. Alan R. Liss, Inc., New York.
11. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization: a method for isolation of cloned cDNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72**:3961–3965.
12. **Houck, C. M., F. P. Rinehart, and C. W. Schmid.** 1979. A ubiquitous family of repeated DNA sequences in the human genome. J. Mol. Biol. **132**:289–306.
13. **Hwu, H. R., J. W. Roberts, E. H. Davidson, and R. J. Britten.** 1986. Insertion and/or deletion of many repeated DNA sequences in human and higher ape evolution. Proc. Natl. Acad. Sci. USA **83**:3875–3879.
14. **Jagadeeswaran, P., B. G. Forget, and S. M. Weissman.** 1981. Short interspersed repetitive DNA elements in eucaryotes: transposable DNA elements generated by reverse transcription of RNA pol III transcripts? Cell **26**:141–142.
15. **Maxam, A. M., and W. Gilbert.** 1976. A new method of sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
16. **Milner, R. J., and J. G. Sutcliffe.** 1983. Gene expression in rat brain. Nucleic Acids Res. **11**:5497–5520.
17. **Penman, S., K. Scherrer, Y. Becker, and J. E. Darnell, Jr.** 1963. Polyribosomes in normal and poliovirus-infected HeLa cells and their relationship to messenger-RNA. Proc. Natl. Acad. Sci. USA **49**:654–662.
18. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3′ non-coding region sequences in eukaryotic messenger RNA. Nature (London) **263**:211–214.
19. **Rogers, J.** 1985. The origin and evolution of retroposons. Int. Rev. Cytol. **93**:187–279.
20. **Schibler, K., M. Tosi, A. C. Pittet, L. Fabiani, and P. K. Wellauer.** 1980. Tissue-specific expression of mouse α-amylase genes. J. Mol. Biol. **142**:93–116.
21. **Schmid, C. W., and W. R. Jelinek.** 1982. The Alu family of dispersed repetitive sequences. Science **216**:1065–1070.
22. **Sutcliffe, J. G., R. J. Milner, F. E. Bloom, and R. A. Lerner.** 1982. Common 82-nucleotide sequence unique to brain RNA. Proc. Natl. Acad. Sci. USA **79**:4942–4946.
23. **Sutcliffe, J. G., R. J. Milner, J. M. Gottesfeld, and R. A. Lerner.** 1984. Identifier sequences are transcribed specifically in brain. Nature (London) **308**:237–241.
24. **Sutcliffe, J. G., R. J. Milner, J. M. Gottesfeld, and W. Reynolds.** 1984. Control of neuronal gene expression. Science **225**: 1308–1315.
25. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5215.
26. **Ullu, E., S. Murphy, and M. Melli.** 1982. Human 7SL RNA consists of a 140 nucleotide middle-repetitive sequence inserted in an Alu sequence. Cell **29**:195–202.
27. **Ullu, E., and C. Tschudi.** 1984. Alu sequences are processed 7SL RNA genes. Nature (London) **312**:171–172.
28. **Ullu, E., and A. M. Weiner.** 1984. Human genes and pseudogenes for the 7SL RNA component of signal recognition particle. EMBO J. **3**:3303–3310.
29. **Ullu, E., and A. M. Weiner.** 1985. Upstream sequences modulate the internal promoter of the human 7SL RNA gene. Nature (London) **318**:371–374.
30. **Van Arsdell, S. W., R. A. Denison, L. B. Bernstein, A. M. Weiner, T. Manser, and R. F. Gesteland.** 1981. Direct repeats flank three small nuclear RNA pseudogenes in the human genome. Cell **26**:11–17.
31. **Weiner, A. M., P. L. Deininger, and A. Efstratiadis.** 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu. Rev. Biochem. **55**:631–661.