Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 11315–11319, December 1991
Medical Sciences

Inactivation of the cholinesterase gene by *Alu* insertion: Possible mechanism for human gene transposition

(retrotransposon/homologous recombination/small polydisperse circular DNAs)

KOJI MURATANI*, TOSHIKAZU HADA*†, YOSHIHIRO YAMAMOTO‡, TADASHI KANEKO§, YOSHIHISA SHIGETO¶, TORU OHUE*, JUNICHI FURUYAMA‡, AND KAZUYA HIGASHINO*

The *Third Department of Internal Medicine and ‡Department of Genetics, Hyogo College of Medicine, Nishinomiya, Hyogo 663, Japan; §Department of Surgery, Kinki Central Hospital of the Mutual Aid Association of Public School Teachers, Itami, Hyogo 664, Japan; and ¶Third Department of Internal Medicine, Osaka University School of Medicine, Fukushima, Osaka 530, Japan

Communicated by Clement L. Markert, September 9, 1991

ABSTRACT **The human cholinesterase (ChE) gene from a patient with acholinesterasemia was cloned and analyzed. By using ChE cDNA as a probe, four independent clones were isolated from a genomic library constructed from the patient's DNA. Sequencing analysis of all of the four clones revealed that exon 2 of the ChE gene was disrupted by a 342-base-pair (bp) insertion of *Alu* element, including a poly(A) tract of 38 bp, which showed 93% sequence homology with a current type of human *Alu* consensus sequence. Southern blot analysis showed that the *Alu* insertion occurred in both alleles of the patient and was inherited in the patient's family. This *Alu* insertion was flanked by 15-bp of target site duplication in exon 2 corresponding to positions 1062–1076 of ChE cDNA, indicating that an *Alu* element could have been integrated by retrotransposition. Thus, this case provides an important clue to the mechanism of inactivation of a gene by integration of a retrotransposon.**

Cholinesterase (ChE; acylcholine acylhydrolase, EC: 3.1.1.8; also known as butyrylcholinesterase) is a ubiquitous, polymorphic family of choline ester-hydrolyzing enzymes. The physiological function of ChE is unknown, but patients with acholinesterasemia have been reported to have prolonged apnea when they undergo surgery with injection of the myorelaxant succinylcholine (1). Thus, as the ChE deficiency has, otherwise, no influence on health, there could not be any selection to eliminate the deficient ChE allele in the population. Therefore, it is easy to find the subjects with ChE deficiency and ChE is thought to be a good model for use in studies on human gene mutation.

In plasma, ChE is mainly present in a water-soluble form of a G_4 tetramer, each subunit consisting of 574 amino acids of the mature protein truncated by 28 amino acids of signal peptide. Full-length cDNA of 2.4 kilobases (kb) has been sequenced in two laboratories (2, 3). Genomic DNA has been found to have three introns at positions corresponding to nucleotides −93, 1433, and 1600 of ChE cDNA (4).

ChE is genetically linked to two independent loci, *CHE1* and *CHE2* (5). Locus *CHE1* is on chromosome 3q21–q26 and expresses almost all of the ChE. Locus *CHE2* is on chromosome 16p11–q23 and encodes a C_5 variant that gives a distinct band on agarose gel electrophoresis in about 10% of Caucasians (5).

Only a point mutation has been reported to change the activity of ChE, including atypical (dibucaine-resistant) and silent (complete loss of enzyme activity) ChE. In atypical ChE, a point mutation was identified at nucleotide 209 (adenine to guanine) that converts Asp-70 to glycine. As Asp-70 is thought to be a component of the anionic binding site, its conversion to glycine reduces the affinity of cholinesterase for choline esters (6). In silent ChE, a frame shift mutation at Gly-117 (GGT to GGAG) was found in the family member with silent ChE (7).

Table 1. Sequences and locations of oligonucleotide primers used for amplification

Primer name	Nucleotide location*	Sequence (5′ → 3′)
AP1	−77 to −54	GCAAAGTCACAATCATATGCATCAG
AP2†	Intron 2	GCTCTGTGAACAGTGTTAGAAAACA
AP3‡	Intron 3	CCGTGCCTTGGAGAGTATACTTCATCC
AP4§	Intron 3	CGAAATTATTTTTCAGTTAATGAAACAG
AP5	718–694	TAGCTTCATAAAGAGATGTTACCGC
AP6	1131–1155	AGATGATCAGAGACCTGAAAACTAC
AP7	1858–1834	GGAATCAATATTATCCTTCTGGCAT

*Position +1 is the first nucleotide of the coding sequence for the mature protein.
†AP2 matches the sequence in intron 2 near the 3′ boundary with exon 3.
‡AP3 is complementary to the template strand near the 5′ end of intron 3.
§AP4 matches the sequence near the 3′ terminus of intron 3.

Several other mechanisms, such as deletion, base addition, and insertion have been reported to inactivate other human genes. Recently, the mutation of certain human genes by large nonviral insertions has been demonstrated. In one study, an L1 insertion into exon 14 of the factor VIII gene was found in two patients with hemophilia A from different families (8). In another study, an experimental *Alu* transposition was demonstrated in a UV-irradiated human lung carcinoma cell line, A549 (9). Since duplication of the target site and a poly(A) sequence at the 3′ end were found in the insertion, these gene families have been regarded as nonviral retrotransposons. Most of the transposable elements found in human DNA have repetitive sequences, and certain classes of LINEs (10) and SINEs (10) (long and short interspersed repetitive elements) are thought to be retrotransposons. However, the exact mechanism of retrotransposition remains unknown.

We describe an *Alu* insertion into the ChE gene in the DNA isolated from a patient with acholinesterasemia, indicating a possible mechanism for retrotransposition involving the *Alu* element.

MATERIALS AND METHODS

DNA Isolation. A 60-yr-old man (II-5) was admitted to our hospital because of diabetes mellitus, and the lack of ChE

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: spc, small polydisperse circular; PCR, polymerase chain reaction; ChE, cholinesterase.
†To whom reprint requests should be addressed.

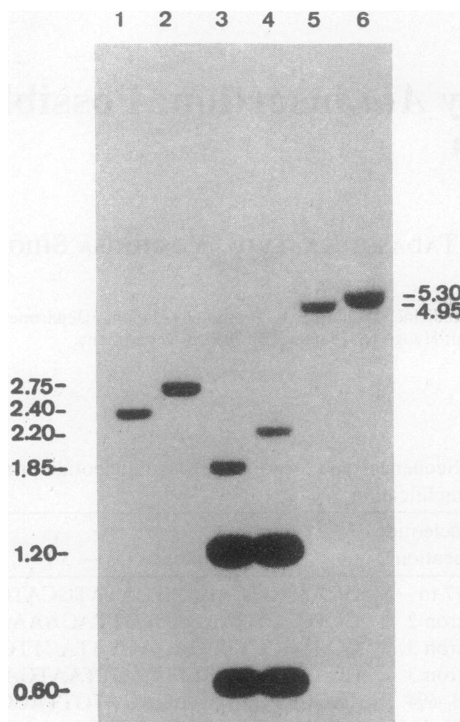

FIG. 1. Southern blot hybridization of the mutant ChE gene. Samples of genomic DNA (10 μg) were obtained from the patient with acholinesterasemia (lanes 2, 4, and 6) and a normal control (lanes 1, 3, and 5). Both samples were subjected to complete enzymatic digestion with restriction endonucleases *Eco*RI (lanes 1 and 2), *Eco*RI/*Bam*HI (lanes 3 and 4), and *Hin*dIII (lanes 5 and 6).

activity in his serum was found by chance. The B lymphoblastoid cell line, Lch21, was established from peripheral mononuclear cells of this patient with acholinesterasemia, II-5, by Epstein–Barr virus infection. DNA was isolated according to the method previously described (11).

DNA Amplification by the Polymerase Chain Reaction (PCR). DNA amplification was carried out by PCR with Taq polymerase (Perkin–Elmer/Cetus) and sets of primers (Table 1). A total of 25–40 cycles of amplification was performed, each cycle consisting of a denaturation step at 94°C for 1 min, an annealing step at 48°C for 2 min, and an extension step at 72°C for 5–7 min.

Construction of the Full-Length cDNA Probe for the ChE Gene. The 795-base-pair (bp) fragment corresponding to positions −77 to 718 was amplified from lymphocyte DNA of a normal subject by PCR with primers AP1 and AP5 and was cut out with *Eco*RI to obtain a 166-bp segment (−77 to 89 of the ChE cDNA). The 1201-bp *Eco*RI/*Bal* I segment (positions 90–1290 of the cDNA) was obtained by cloning a normal genomic DNA library. The 728-bp fragment (1131–1858 of cDNA) was amplified from first-strand cDNA of HuH7 (Human Hepatoma cell line) (12) by PCR with primers AP6 and AP7 (Table 1). This fragment was cut out with *Bal* I, resulting in a 568-bp segment (1291–1858 of cDNA). Each fragment was ligated with the adjacent fragment to construct the full-length ChE cDNA. The resulting full-length cDNA, named pc1935, was confirmed to have an identical sequence to the ChE cDNA reported previously (2, 3).

Southern Blot Hybridization. Southern blot analysis was done with a full-length cDNA, pc1935, as a probe by the method described previously (13).

Construction of a Genomic Library and Screening. The genomic DNA isolated from the B lymphoblastoid cell line Lch21 was digested with *Eco*RI. The digest was separated by electrophoresis in 0.8% low-melting-point agarose gel, and the DNA fragments of 2.5–3.0 kb were isolated and ligated into the λgt10 arms with T4 DNA ligase. The ligated material was packaged by using Gigapack gold (Stratagene) and titered on NM514 cells to yield 8 × 10^5 independent phage plaques. The 4 × 10^5 recombinants were screened without amplification. After transfer to nylon membrane filters, the plaques were lysed and hybridized with ^{32}P-labeled ChE cDNA at 42°C according to the manufacturer's instructions.

DNA Amplification of the Remaining Segment of the Patient's ChE Gene. PCR was used for analyzing the genomic sequences other than the abnormal 2.75-kb *Eco*RI fragment. The 795-bp segment in exon 2, corresponding to positions −77 to 718 of cDNA, was amplified by PCR to examine the region upstream of the *Eco*RI site (nucleotide 90), with the primers AP1 and AP5 (Table 1). Then the 336-bp segment containing all of exon 3 was amplified with primers AP2 and AP3 (Table 1). The segment of 327 bp containing all of the coding region of exon 4 was amplified with primers AP4 and AP7 (Table 1). All of these fragments were blunted at their ends by T4 DNA polymerase and subcloned in pUC18.

DNA Sequencing. DNA fragments were subcloned in pUC18 and sequenced by the dideoxy chain-termination method (14) with the sequence primers M4 and RV (Takara Shuzo, Kyoto).

RESULTS

Southern Blot Hybridization Analysis. Examination of serum ChE activity of the patient II-5 showed the absence of butyrylcholine hydrolyzing activity. The possibility that

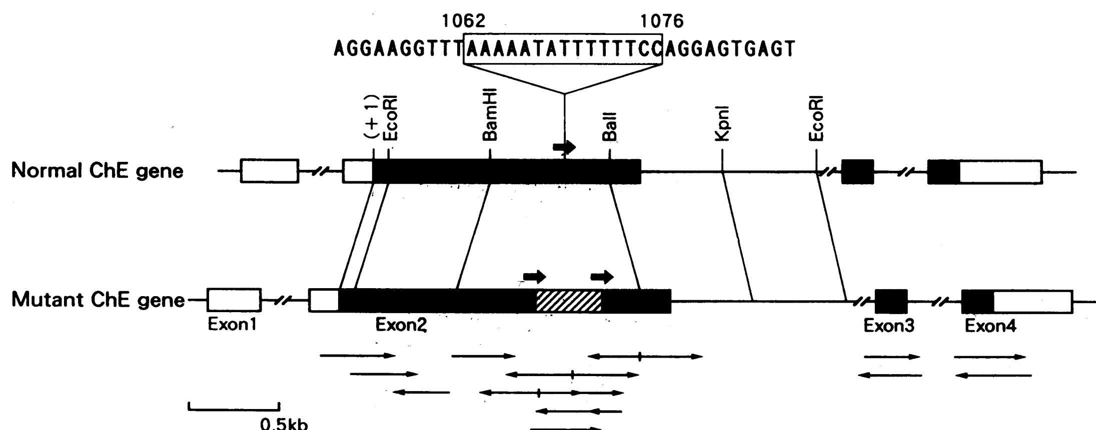

FIG. 2. Restriction map and sequencing strategy of the mutant ChE gene. The solid bar represents the coding sequence for the mature protein of normal CHE. The hatched bar represents the inserted *Alu* element. Broad arrows and boxed sequence indicate the target site of insertion, and narrow arrows indicate the directions and extents of DNA sequencing.

```
Mutant ChE  [AAAGAATTTCAGGAAGGTTT]AAAAATATTTTTCCGGCCGGGCGCGGTGGCTCACGCCTGTA
Cons. Alu                       --------------->  ...........................
                                 direct repeat    stop

Mutant ChE  ATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCATGAGGTCAGGAGATCGAGACCATCCT
Cons. Alu   ............................C......C.........................

Mutant ChE  GGCTAACAAGGTGAAACCCCGTCTCTACTAAAAAATACAAAAAATTAGCCGGGCGCGGTGGC
Cons. Alu   ........C..........................*...........................T......

Mutant ChE  GGGCGCCTGTAGTCCCAGCTACTGGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAAG
Cons. Alu   .......................C......................................G.

Mutant ChE  CGGACGTTGCAGTGAGCCGAGATTGTGCCACTGCAGTCCGCAGTCCGGCCTGGGCGACAGAG
Cons. Alu   ....GC..................C.C.........C...**..*..****............

Mutant ChE  CGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATTT
Cons. Alu   ..............                                          ___________
                                                                   direct repeat
Mutant ChE  TTTCC[AGGAGTGAGTGAGTTTGGAA]
            ----->
```

FIG. 3. Nucleotide sequence of inserted *Alu* element. Cons. Alu represents the sequence of the human *Alu* consensus sequence (15). Boxes enclose the sequence of *CHE*. Arrows indicate 15 bp of direct repeat. Dots and asterisks indicate the same nucleotides and the nucleotides that are not present in the *Alu* consensus sequence, respectively.

expression of the silent ChE phenotype was due to rearrangement of the ChE gene was examined by Southern blot hybridization with the full-length cDNA probe, pc1935. The patient's DNA gave the 2.75-kb fragment for *Eco*RI but not a band at the 2.4-kb fragment obtained with the normal ChE gene (Fig. 1). Hence, there was a 0.35-kb elongation in both alleles of the patient's DNA.

Analysis of the Patient's ChE Gene. To identify the 0.35-kb elongation of the *Eco*RI fragment in the patient's ChE gene, 4×10^5 clones of an unamplified genomic DNA library obtained from the patient's DNA were screened with ChE cDNA as a probe, and four independent clones (mc1, mc2, mc3, and mc4) were obtained, all of which contained 2.75-kb fragment. These 2.75-kb *Eco*RI fragments from mc1–4 clones were recloned into pUC18.

The restriction maps of mc1–4 revealed that the four clones were identical, and a 0.35-kb fragment was inserted in exon 2 (Fig. 2). For determination of the nucleotide sequence of the inserted element and the 5′ and 3′ boundaries, the 2.75-kb *Eco*RI fragment was analyzed.

The inserted element was 342 bp long and was flanked by 15 bp of target site duplication in the boundary positions 1062–1076 corresponding to normal ChE cDNA. The sequence of this inserted element showed 93% homology with the human *Alu* consensus sequence of the evolutionarily most recent branch (i.e., the class IV of Britten *et al.*) (15) (Fig. 3). In addition to the data of Southern blot analysis, the same *Alu* element was inserted into the same position in all mc1–4 clones, thus indicating the *Alu* element was inserted into both *CHE1* and *CHE2* alleles homozygously. PCR amplification was used to examine the region upstream of the *Eco*RI site (nucleotide 90) of exons 2, 3, and 4. On sequencing the PCR products, no other mutation was found except the *Alu* insertion in exon 2.

Family Study. The parents of patient II-5 were dead, but three siblings and two children are alive (Fig. 4 *Upper*). Enzymatic studies revealed that the patients II-5 and II-1 had no serum ChE activity and that III-1 and III-2 had half the normal level of ChE activity. Southern blot analysis showed that DNA from II-1 gave only the same *Eco*RI band of 2.75 kb as that of patient II-5 and that the DNA of III-1 and III-2 gave *Eco*RI bands of both 2.75 kb and 2.4 kb (Fig. 4 *Lower*). These data indicate that the *Alu* insertion was inherited in this family. As the parents of the patient were not consanguineous, the ancestral mutation in the ChE gene may be conserved regionally, resulting in expression of a silent ChE phenotype in a homozygous individual.

DISCUSSION

An insertion of an *Alu* element into the ChE gene was identified in the patient with acholinesterasemia. Immunoassays for the ChE protein were performed by Ouchterlony's double immunodiffusion assay with polyclonal antibody to

FIG. 4. (*Upper*) Family tree of the patient with acholinesterasemia. Boxes indicate males and circles indicate females. The arrow indicates the patient studied (II-5). Homozygous silent individuals are represented by solid figures, and heterozygotes are represented by half-open/half-solid figures. Other persons indicated by open figures were not studied. (*Lower*) Southern blot analysis of family members. Samples of genomic DNA (10 μg) were digested completely with *Eco*RI. Lanes: 1, normal control; 2, patient II-5; 3, son III-1; 4, daughter III-2; 5, brother II-1.

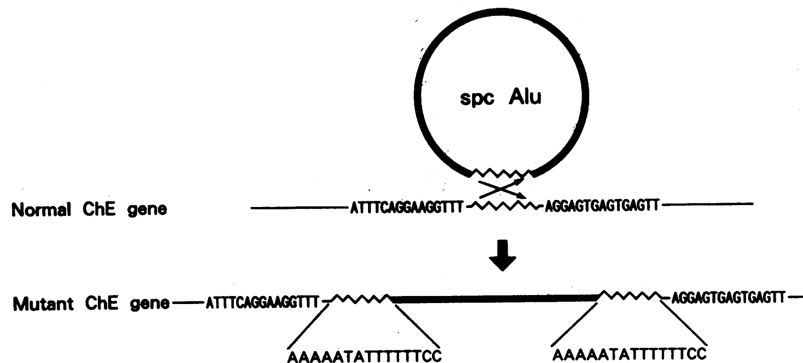

FIG. 5. Schematic representation of *Alu* integration in the form of spc DNA. The solid bar represents the *Alu* element. The wavy line represents the target site of the ChE gene and the identical region of the spc DNA including the *Alu* element.

ChE and by Sandwich immunoassay with a monoclonal antibody against ChE, although the epitope of monoclonal antibody has not been determined. Both assays could not detect the immunoreactive protein of ChE at all (data not shown). Two possibilities exist for the loss of ChE protein, either (*i*) normal transcription was interrupted by the inserted *Alu* element or (*ii*) mature mRNA was transcribed, but the translation was stopped at the stop codon in the inserted *Alu* element, resulting in the loss of the stability of the truncated protein. We could not further evaluate the different hypotheses, since the tissues (brain or liver) expressing the mRNA of ChE could not be obtained from the patient.

In this paper, we identified an insertion of an *Alu* element into the ChE gene. *Alu* elements are a specific human family of interspersed repetitive sequences, with over 100,000 copies per genome, and their dispersion and high copy number have been thought to be due to transpositions. We found that the inserted *Alu* element created a 15-bp direct repeat at both junctions with no mutation in the original ChE gene sequence. These data indicate that the *Alu* insertion might not be due to homologous recombination between linear DNA fragments, but to integration by retrotransposition.

One mechanism has been proposed that an *Alu* element is integrated into a genomic site with staggered nicks (16). Nicking of the genome on both strands followed by repair system could make the flanking direct repeats at the integration site, but there are several problems to be solved. First, the target site must be nicked to allow the entry of the new sequence, but the variation in length and nucleotide sequence of the flanking direct repeat suggest that the specific integration enzymes, as in the case for retroviruses and transposons, could not be used. Alternatively random nicking from chemicals or radiations could be involved to make staggered nicks, but preferential adenosine richness at the flanking direct repeat rather reflects the selection of specific sequence during the integration process. Second, the reverse transcribed *Alu* element (cDNA of *Alu* element) must be integrated into the genomic site with staggered nicks. It is hypothesized that an adenosine richness at the nick site may interact with the thymidine-rich cDNA end [reverse-transcribed poly(A) tract of the *Alu* element] to stabilize the interaction at the 5′ end of the cDNA. But unknown enzymes or mechanisms are needed for the ligation of the 3′ end of the cDNA to the target site. Thus, the mechanism of retrotransposition is not fully elucidated.

We speculate another possible mechanism that an *Alu* element was integrated into a new site of the genome. In yeast, it has been found that a circular plasmid carrying a yeast gene was integrated into a homologous region on the genomic DNA, resulting in sequence duplication of the target site. This phenomenon has been demonstrated in several genes of yeast (17–19). Both yeast and humans are eukaryotes. They have a similar gene control system, and the same phenomenon could occur when a circular form of DNA is present in human cells.

Recently small polydisperse circular (spc) DNAs have been shown in many types of human cells by electron microscopy. Kunisada and Yamagishi (20) reported that repetitive sequences including *Alu* elements were found in most of the spc DNAs of HeLa cells. Furthermore Krolewski and Rush (21) cloned *Alu*-containing spc DNAs after isolation of 300-bp spc DNAs. Thus, we propose that the spc DNA, including *Alu* element, could be integrated into the ChE gene by homologous recombination. The spc DNA containing both the *Alu* element and a homologous sequence to the insertional boundary of ChE has not been found, and this casts a question on the origin of the homologous sequence to the ChE. Although termination of RNA polymerase III needs four or more thymidine residues, *Alu* consensus sequence does not contain such a termination codon. Therefore, a termination codon in a sequence adjacent to the 3′ end of the *Alu* element must be used. Furthermore, reverse transcripts of *Alu*-specific mRNAs are thought to be possible precursors for the generation of spc DNA.

Thus, we suggest that transposable *Alu* elements and adjacent sequence to the termination codon of RNA polymerase III left their parental genome through the RNA intermediate, converted to the form of spc DNA with some kind of reverse transcriptase, and could be integrated into a new site of a chromosome by homologous recombination (Fig. 5). We cannot conclude at present which hypothesis is correct, integration of an *Alu* element with staggered nicks or integration by homologous recombination between the ChE and the spc DNA including the *Alu* element. To clarify the mechanism of *Alu* retrotransposition, more critical data that reproduce some parts of the retrotransposition process *in vitro* would have to be gathered.

This work was supported in part by a grant from the Japan Foundation for Applied Enzymology and a grant-in-aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

1. Kalow, W. & Davies, R. O. (1958) *Biochem. Pharmacol.* **1**, 183–192.
2. Prody, C. A., Zevin-Sonkin, D., Gnatt, A., Goldberg, O. & Soreq, H. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3555–3559.
3. McTiernan, C., Adkins, S., Chatonnet, A., Vaughan, T. A., Bartels, C. F., Kott, M., Rosenberry, T. L., La Du, B. N. & Lockridge, O. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6682–6686.
4. Arpagaus, M., Kott, M., Vatsis, K. P., Bartels, C. F., La Du, B. N. & Lockridge, O. (1990) *Biochemistry* **29**, 124–131.
5. Soreq, H., Zamir, R., Zevin-Sonkin, D. & Zakut, H. (1987) *Hum. Genet.* **77**, 325–328.
6. McGuire, M. C., Nogueira, C. P., Bartels, C. F., Lightstone,

H., Hajra, A., Van Der Spek, A. F. L., Lockridge, O. & La Du, B. N. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 953–957.
7. Nogueira, C. P., McGuire, M. C., Graeser, C., Bartels, C. F., Arpagaus, M., Van der Spek, A. F. L., Lightstone, H., Lockridge, O. & La Du, B. N. (1990) *Am. J. Hum. Genet.* **46,** 934–942.
8. Kazazian, H. H., Wong, C., Youssoufian, H., Scott, A. F., Phillips, D. G. & Antonarakis, S. E. (1988) *Nature (London)* **332,** 164–166.
9. Lin, C. S., Goldthwait, D. A. & Samols, D. (1988) *Cell* **54,** 153–159.
10. Singer, M. F. (1982) *Cell* **28,** 433–434.
11. Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3,** 2303–2308.
12. Nakabayashi, H., Taketa, K., Miyano, K., Yamane, T. & Sato, J. (1982) *Cancer Res.* **47,** 3858–3863.
13. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
14. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143,** 161–178.
15. Britten, R. J., Baron, W. F., Stout, D. B. & Davidson, E. H. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4770–4774.
16. Deininger, P. L. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington), pp. 619–636.
17. Botstein, D. & Davis, R. W. (1982) in *The Molecular Biology of the Yeast Saccharomyces, Metabolism and Gene Expression*, eds. Strathern, J. N., Jones, E. W. & Broach, J. R. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 607–638.
18. Scherer, S. & Davis, R. W. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4951–4955.
19. Shortle, D., Harber, J. & Botstein, D. (1982) *Science* **217,** 371–373.
20. Kunisada, T. & Yamagishi, T. (1987) *J. Mol. Biol.* **198,** 557–565.
21. Krolewski, J. J. & Rush, M. G. (1984) *J. Mol. Biol.* **174,** 31–40.