Nucleic Acids Research, Vol. 19, No. 13 3619–3623

Amplification dynamics of human-specific (HS) alu family members

Mark A.Batzer[1], Vandana A.Gudi[2], J.Carlos Mena[1], David W.Foltz[3], Rene J.Herrera[2] and Prescott L.Deininger[1,4]

[1]Department of Biochemistry and Molecular Biology, Louisiana State University Medical Center, 1901 Perdido Street, New Orleans, LA 70112, [2]Department of Biological Sciences, Florida International University, University Park Campus, Miami, FL 33199, [3]Department of Zoology and Physiology, Louisiana State University, Baton Rouge, LA 70803 and [4]Laboratory of Molecular Genetics, Alton Ochsner Medical Foundation, New Orleans, LA 70121, USA

Received March 19, 1991; Revised and Accepted April 26, 1991

ABSTRACT

We have investigated the distribution of several recently inserted Alu family members within representatives of diverse human groups. Human population studies using 65 unrelated human DNA samples, as well as a familial study to test inheritance, showed that individual Alu family members could be divided into three groups. The first group consisted of relatively older Alu family members which were monomorphic (homozygous) throughout the population tested (HS C3N1 and C4N6). The second group (HS C4N2, C4N5 and C4N8), apparently inserted into other repetitive regions of the genome, resulting in inconclusive results in the PCR test used. However, it is clear that these particular Alu insertions were present in a majority if not all of the loci tested. The third group was comprised of three dimorphic Alu family members (HS C2N4, C4N4 and TPA 25). Only a single Alu family member (TPA 25) displayed a high degree of dimorphism within the human population. This latter example also showed different allele frequencies in different human groups. The isolation and characterization of additional highly dimorphic Alu family members should provide a useful tool for human population genetics.

INTRODUCTION

The Alu family of short interspersed repetitive DNA elements (SINEs) is distributed throughout primate genomes (for reviews see 1, 2, 3). The Alu family represents one of the most successful classes of mobile elements, having arisen as a repetitive DNA family within the last sixty-five million years (4) and amplified to a copy number in excess of 500,000 within the human genome (5). Each Alu element is approximately 300 bp in length, consisting of two tandemly arranged halves, with the right half containing an additional 31 bp relative to the left half (5). Alu elements also contain a middle A-rich region, a 3′ oligo-dA tail which is variable in length, and are flanked by short direct repeats which form during integration at staggered chromosomal nicks (5, 6). Alu sequences are thought to be ancestrally derived from the 7SL RNA gene (7), and mobilize through an RNA polymerase III derived intermediate via a process termed retroposition (8).

The Alu family members found within primate genomes may be subdivided into groups of related subfamily members based on nucleotide identity (9). Several overlapping subfamilies of various genetic ages have been identified within primate genomes (10, 11, 12, 13, 14). The most recently formed subfamily of Alu sequences found within the human genome was originally referred to as the 'new' subfamily (15). It has subsequently been further characterized and termed the Predicted Variant (PV) subfamily (16) and the Human-Specific (HS) subfamily (17). Interestingly, the HS subfamily (PV) was also found to be transcriptionally active *in vivo,* suggesting the transpositional competence of this subfamily (16). There are an estimated 500 (17) to 2000 (16) recently inserted HS subfamily members in the human genome. Individual HS subfamily members share a high degree of nucleotide identity (>98%) with the subfamily consensus sequence suggesting that they were derived from a single, or at most a closely related set of, source gene(s) (6, 15, 16, 17, 18). Previously, a number of members of the HS subfamily were found to be present only within the human genome, and absent at orthologous positions in the genomes of other primates, indicating that most, if not all, of the HS subfamily members had amplified within the human genome in the last 4–6 million years (17). In this report, we present a detailed analysis of the distribution of eight recently-inserted HS subfamily members within major human groups.

MATERIALS AND METHODS

DNA Samples and Cell Lines

Individual DNA samples were isolated from peripheral human lymphocytes as previously described (19). These samples were made available from previous studies and considered exempt from human subjects restrictions following IRB review. African DNA samples were provided by Dr. Haig H. Kazazian. Familial DNA

samples were a gift from Dr. Bronya Keats. Human-hamster hybrid cell line 'PCRable' DNA samples were obtained from BIOS.

PCR Amplification

Amplification of DNA samples was carried out in 100 ml reactions using 100 ng of target DNA, 750 ng of each primer, 100 mM dNTP's and *Taq* DNA polymerase (3 u) according to the supplier's (BIOS) instructions. Each sample was subjected to the following amplification cycle; 1 min at 94°C (denature), 2 min at the annealing temperature, and 2 min at 72°C (extension), for 30 cycles. Oligonucleotide primers and annealing conditions for HS C2N4, C3N1, C4N4, C4N6, C4N8, C4N5 and TPA 25 were previously reported (17). Subfamily members HS C2N4, C3N1, C4N2, C4N4, C4N5, C4N6 and C4N8 were isolated from a random human genomic library as described previously (17). The TPA 25 Alu family member was found in one of two independent clones of the human tissue plasminogen activator locus (20). The oligonucleotide primers for HS C4N2 were 5′ primer 5′-TGCAGGAATTCAGCACAAATTGTAG-3′ and 3′ primer 5′-AAATCAGTCCTACCATGATTTTGTC-3′ with an annealing temperature of 58°C. Twenty microliters of each sample was fractionated on a 2% agarose gel with 0.5 μg/ml ethidium bromide. Reaction products were directly visualized by UV fluorescence.

Computer Analyses

Allele frequencies and chi-square analyses were determined using a genetic data analysis program designed for SAS (21). Individual Alu family member pre-integration sites (HS C4N2, C4N5 and C4N8) were used to search the EMBL data base (release 21) using the FASTNSCAN program of PC/GENE (Intelligenetics) with a k-tuple value of 4. This program is based on the FASTN algorithm of Lipman and Pearson (22).

RESULTS

Gross Distribution of Recently Amplified Alu Family members

Each Alu HS subfamily member insertion locus was assayed by PCR analysis using oligonucleotide primers complementary to the unique DNA sequences flanking each Alu subfamily member. These primers amplified a DNA fragment approximately 400 bp long if the Alu sequence was present, or approximately 150 bp in the absence of the Alu family member (see Figure 1). Individuals were scored as either homozygous (+ +), heterozygous (+ −), or homozygous (− −) for the presence of the 400 bp band (+), or the 150 bp band (−). The genotypes, expected values and allele frequencies of the individual HS subfamily members determined from PCR analyses are summarized in Table 1. Individual subfamily members were distributed in three distinct groups based upon their distribution within the human population. One group was comprised of dimorphic subfamily members TPA-25, HS C2N4 and HS C4N4. A second group (HS C3N1 and C4N6) were monomorphic (homozygous) for the presence of an Alu family member in the 65 unrelated human samples which were analyzed (Table 1). The third group (HS C4N2, C4N5 and C4N8) displayed both bands in all individuals tested, apparently as the result of insertion of Alu subfamily members into other repetitive loci (see below).

The dimorphic subfamily members may be further subdivided into highly dimorphic (TPA 25) and weakly dimorphic (HS C2N4 and C4N4) groups. The presence of only one and two heterozygous individuals for HS C2N4 and C4N4, respectively (Table 1), make these Alu family members uninformative at the population level, but are consistent with the relatively recent insertion of the subfamily in the human genome. Only the TPA 25 subfamily member showed significant heterozygosity within groups using a Chi-square test for homogeneity on allele counts ($\chi^2=12.05$, d.f.$=3$, $P<0.01$). The allelic frequency of the TPA 25 subfamily member was significantly different in populations of African origin as compared to those of non-African origin ($\chi^2=9.93$, d.f.$=1$, $P<0.01$). In all instances, genotypic frequencies within groups were close ($P<0.05$) to the predicted (Hardy-Weinberg) frequencies using Chi-square tests for goodness-of-fit (Table 1).

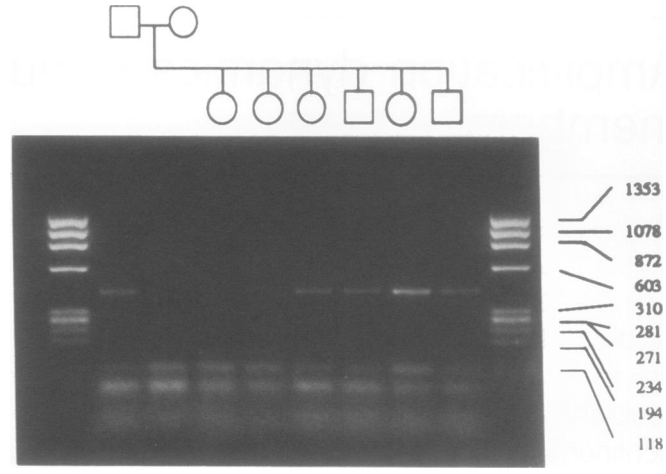

Figure 1. Inheritance of the TPA 25 subfamily member. Pedigree of a family assayed for the transmission of the TPA 25 subfamily member and an agarose gel of the PCR products from the amplification of the TPA 25 subfamily member. The fragment length markers (outer lanes, Hae III digested ΦX174 RF DNA) are indicated in bp. Experimental lanes 1 and 2 represent parents, with lanes 3−8 representing progeny as indicated. The lower cloud and band were also present in control experiments containing no genomic DNA and represent excess primer, nucleotides and primer dimer amplification respectively. Parental lane 1 represents a heterozygote, with both the 150 and 400 bp bands. Parental lane 2 represents homozygosity for the 150 bp band. The progeny lanes all show one of these two patterns.

Pedigree Analyses

We performed PCR analysis on three sets of familial DNA samples comprised of the parents and a total of 15 progeny to ascertain the mode of inheritance for some of the recently amplified Alu family members (TPA 25, HS C4N2, C4N5 and C4N8). The TPA 25 subfamily member demonstrated an inheritance pattern consistent with normal Mendelian codominant inheritance (Figure 1). The other three showed both the large and small PCR bands in every sample tested. A representative pedigree and chromatograph of the resultant PCR products from the HS C4N5 locus is shown in Figure 2. These data show that the HS C4N5 subfamily member from a representative family shows both bands in every individual tested. A Chi-square analysis showed that the observed ratio of progeny (0:1:0) was significantly different from that expected (1:2:1) from a cross of two heterozygous individuals at a single locus ($\chi^2=15$, d.f.$=2$, $P<0.001$). Subfamily members HS C4N2 and C4N8 also displayed both bands in every individual tested (Table 1).

Table 1. Distribution of Recently Amplified Alu Family Members[1,2]

		Subfamily member							
		TPA 25	HS C2N4	HS C4N4	HS C3N1	HS C4N6	HS C4N2	HS C4N5	HS C4N8
Asians	+ +	10 (8.22) {0.66}	19 (19) {1.0}	18 (18) {0.97}	19 (19) {1.0}	19 (19) {1.0}	0 (4.75) {0.5}	0 (4.75) {0.5}	0 (4.75) {0.5}
	+ −	5 (8.56)	0 (0)	1 (0.97)	0 (0)	0 (0)	19 (9.5)	19 (9.5)	19 (9.5)
	− −	4 (2.22) {0.34}	0 (0) {0}	0 (0.03) {0.03}	0 (0) {0}	0 (0) {0}	0 (4.75) {0.5}	0 (4.75) {0.5}	0 (4.75) {0.5}
Caucasians	+ +	10 (9.14) {0.63}	22 (22) {1.0}	22 (22) {1.0}	22 (22) {1.0}	22 (22) {1.0}	0 (5.5) {0.5}	0 (5.5) {0.5}	0 (5.5) {0.5}
	+ −	9 (10.72)	0 (0)	0 (0)	0 (0)	0 (0)	22 (11)	22 (11)	22 (11)
	− −	4 (3.14) {0.37}	0 (0) {0}	0 (0) {0}	0 (0) {0}	0 (0) {0}	0 (5.5) {0.5}	0 (5.5) {0.5}	0 (5.5) {0.5}
American Blacks	+ +	4 (3.61) {0.38}	23 (23.01) {0.98}	23 (23.01) {0.98}	24 (24) {1.0}	24 (24) {1.0}	0 (6) {0.5}	0 (6) {0.5}	0 (6) {0.5}
	+ −	11 (11.78)	1 (0.98)	1 (0.98)	0 (0)	0 (0)	24 (12)	24 (12)	24 (12)
	− −	10 (9.61) {0.62}	0 (0.01) {0.02}	0 (0.01) {0.02}	0 (0) {0}	0 (0) {0}	0 (6) {0.5}	0 (6) {0.5}	0 (6) {0.5}
African Blacks	+ +	1 (2.08) {0.42}	NT	NT	NT	NT	NT	NT	NT
	+ −	8 (5.84)							
	− −	3 (4.08) {0.58}							

[1]Genotypes, followed by expected numbers in ( ) and allele frequency in { }.
[2]NT = Not Tested

Pre-integration Site Analyses

The predicted pre-integration sites (the sequences flanking the Alu family members) for HS C4N2, C4N5 and C4N8 were then searched against the EMBL data base (release 21) to identify any significant nucleotide identity with previously analyzed regions of the human genome. Subfamily members HS C4N2 and C4N5 displayed no significant sequence similarity to any previously analyzed regions of the genome. However, HS C4N8 shared 73.7% nucleotide identity with a human Line 1 (L1) repetitive element from the intergenic region of the ϵ and g-γ globin genes (EMBL sequence HSHBEG). Therefore, we believe that HS C4N8 inserted into an L1 family member within the genome. Insertion into a highly repetitive region of the genome, such as an L1 sequence (approximately 100,000 copies) (23), would explain the presence of both bands in every PCR assay. The 150 bp band (no Alu family member) would result from any number of the 100,000 L1 copies (loci) which have not diverged enough to render the oligonucleotide primers ineffective, whereas the 400 bp band results from the amplification of a single Alu family member which is relatively monomorphic throughout the population. However, the true allele frequency of the Alu family member insertion (retroposition) locus (within a single L1 element) remains unknown due to the limitations of the PCR technique. Because the 400 bp band is present in all samples tested, it is clear that the 400 bp allele has a very high frequency, but we cannot directly determine which, if any, of the samples may be heterozygous for this specific allele. However, we can put an upper 95% confidence limit on the frequency of rare, undetected homozygotes for the 150 bp band in the 71 unrelated individuals, which is $1-0.05^{1/71} = 0.041$. The expected frequency of the 150 bp band corresponding to this upper limit is $0.041^{1/2} = 0.203$.

Hybrid Cell Line Analysis

The previous data suggested that all of the HS subfamily members which show both PCR bands in all individuals tested (HS C4N2, C4N5 and C4N8) represented integrations into pre-existing repeated DNA sequences. To further confirm this for HS C4N2, we also amplified a panel of human-hamster hybrid cell line DNAs to determine the distribution of sequences related to the HS C4N2 pre-integration (target) site. Amplification of the panel (data not shown) showed that a 150 bp band (no Alu family member) was amplified from each cell hybrid line of the panel, demonstrating

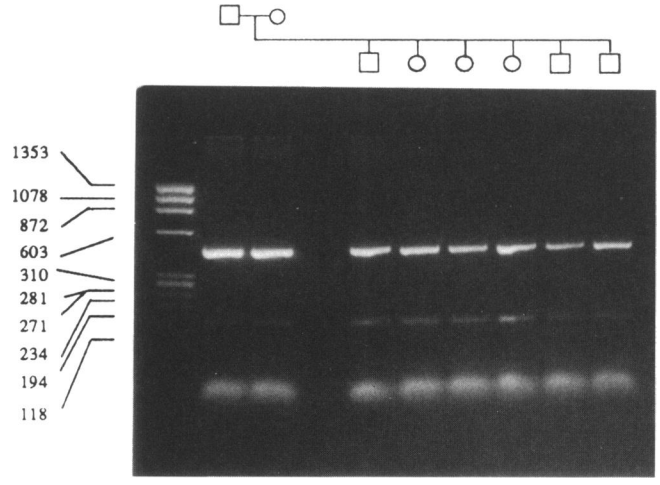

Figure 2. Inheritance of the HS C4N5 subfamily member. Pedigree of a family assayed for the transmission of the HS C4N5 subfamily member and an agarose gel of the PCR products from the amplification of the HS C4N5 subfamily member. The fragment length markers (Hae III digested FX174 RF DNA) are indicated in bp. Lanes 1 and 2 represent parents, with lanes 4−9 representing progeny as indicated in the pedigree.

that the pre-integration site is a repeated sequence located on many chromosomes. The 400 bp band, containing the HS subfamily member, was only located within hybrids containing chromosome 13, 14 or both chromosomes, with the exception of hybrid cell line 968 which contains chromosome 13 and has only the 150 bp band. This latter result makes a definitive chromosomal assignment of the HS C4N2 subfamily member impossible. It may be due to a previously undetected deletion or rearrangement within one of the hybrid cell lines, or may represent dimorphism in the presence of this HS sequence in a human chromosome. However, taken together with the amplification of 65 unrelated individuals and pedigree analysis in which every individual amplified both bands these data clearly demonstrate that HS C4N2 retroposed into a repetitive region of the genome which was dispersed on a number of human chromosomes prior to the insertion of this Alu family member. Therefore we conclude that the HS Alu insertion occurred in a previously uncharacterized repeated sequence, which interferes with the PCR analysis. However, like

HS C4N8 and C4N5, the 400 bp allele must be present at a very high frequency in order to see no individuals who are homozygous for its absence (Table 1).

DISCUSSION

The high percentage (3 out of 11, see below) of recently inserted HS subfamily members which have inserted into duplicated regions of the genome suggests that these types of events occur at a relatively high frequency. Although Alu family members are distributed throughout the genome, they have been shown to cluster within specific regions, for example the introns of the human thymidine kinase (tk), β-tubulin and C-1 inhibitor loci (24, 25 and 26, respectively). However, little more than an A+T rich target site preference for Alu family member insertions has previously been shown (6, 27, 28). The clustering of Alu family members may result from random insertion into additional A+T rich target sites (middle A rich region and 3′ oligo-dA tail) created by previous Alu family member retropositions, some more general insertion site specificity, such as specific chromosomal domains, or selective pressure against insertion into protein encoding regions of the genome. The insertion of such a high proportion (3/11) of recently amplified Alu family members into duplicated regions of the genome suggests that there may be a general chromosomal feature that favors retroposon insertions and/or other sequence duplications. Once a sequence inserts or duplicates in one of these regions, the preference must still exist to favor a second insertion within the same general region.

These studies demonstrate the relative ages of each HS subfamily member. The monomorphic subfamily members (HS C3N1 and C4N6) comprise the oldest group of HS subfamily members. The monomorphic nature of these subfamily members suggests that they predated the origin of modern man 200,000−1 million years ago (29). However, these subfamily members have previously been shown to postdate the human/great ape divergence (17), which is thought to have occurred 4−6 million years ago (30). Studies based upon diagnostic nucleotide substitutions also suggest that the average age of HS subfamily members is 2.8 million years old (6). Therefore, we conclude that these subfamily members retroposed into the human genome 200,000 to 6 million years ago, after the human/great ape divergence and before the radiation of modern man.

The low frequency of Alu family member retropositions (approximately 100−200/million years) (17, 31) suggests that individual Alu family member insertions (retropositions) represent unique, independent insertion events which have occurred only once within the human population. Previous studies, of the primate globin gene clusters, have shown that individual Alu family members remain within primate genomes after their initial insertion and are not subject to significant levels of gene conversion or sequence removal thereby representing stable genomic markers (32, 33, 34). This distinguishes Alu family member insertion dimorphisms from other genomic polymorphisms such as RFLP (35), VNTR (36) and AluVpA (37) which may be of multiple origin and less stable.

Given the unique nature and stability of Alu family member insertions (33), the highly dimorphic state of the TPA 25 insertion within various human groups suggests that this subfamily member is of very recent origin. Previously, the TPA 25 subfamily member was shown to post-date the human/great ape divergence which is thought to have occurred 4−6 million years ago (30) and classified as an HS-2 (6, 17) or PV (18) sub-subfamily member. These studies show that it is present at only about half of the alleles, suggesting that it inserted after the formation of modern man. Studies on mitochondrial DNA have suggested that modern man originated in Africa 200,000−1 million years ago (29), implying that the TPA 25 Alu family member inserted in the human genome in the last 200,000−1 million years. Previously, three other HS subfamily members (PV 92, MLVI-2 and APO) have also been classified as dimorphic within the human genome. PV 92 displayed a high degree of dimorphism within a limited number of individuals (16), while MLVI-2 was present only in one of 59 unrelated individuals (38). The APO HS subfamily member (18), located near the human apolipoprotein AI-CIII-AIV gene cluster, was analyzed by RFLP analysis (39, 40), and was found to have a moderate degree of heterozygosity. Therefore, 55% (6/11) of the randomly selected HS Alu family members appear to be dimorphic, with 18% (2/11) monomorphic and the remaining 27% (3/11) having inserted in duplicated regions of the human genome. Approximately 27% of the HS subfamily members (approximately 135 out of 500 HS subfamily members) may be informative (highly dimorphic) loci for human population studies. The TPA 25 subfamily member allele frequency was significantly different in human groups of African and non-African origins. These differences may reflect that this insertion occurred near the time of migration of modern man out of Africa, or simply reflect genetic drift in these populations having fixed the Alu insertion allele at different frequencies. A single locus on this relatively limited population size does not supply enough resolution to differentiate between the different possibilities. The isolation and characterization of additional highly dimorphic Alu family members in well-defined human population groups should provide a useful tool for understanding the early evolution of modern man and the population dynamics which occurred.

ACKNOWLEDGEMENTS

We would like to thank Drs. V.K.Slagel, G.R.Daniels, and Bronya Keats for critical review of this manuscript. Dr. J.Craig Cohen and Ms. Jane Thompson for providing us with numerous biological samples. We would also like to thank Mr. James Carlton of the LSUMC core laboratories for oligonucleotide synthesis. Photographic services were provided by LSUMC department of photographic services. J.C.M. was supported by the LSU medical student summer research program. This research was supported by USPHS grant number R01 HG00340 to P.L.D. and RR 08205 to R.J.H., and a grant from the Cancer Crusaders to M.A.B.

REFERENCES

1. Schmid, C.W., and Shen, C.-K. J. (1986) In MacIntyre, R.J. (ed.) Molecular Evolutionary Genetics. Plenum Press, N.Y., N.Y., pp. 323−358.
2. Weiner, A.M., Deininger, P.L., and Efstratiadis, A. (1986) *Annu. Rev. Biochem.*, **55**, 631−661.
3. Deininger, P.L. (1989) In Howe, M. and Berg, D. (eds.) Mobile DNA. ASM Press, Washington, D.C., pp. 619−636.
4. Deininger, P.L., and Daniels, G.R. (1986) *Trends Genet.*, **2**, 76−80.
5. Deininger, P.L., Jolly, D.J., Rubin, C.M., Friedmann, T., and Schmid, C.W. (1981) *J. Mol. Biol.*, **151**, 17−33.
6. Batzer, M.A., Kilroy, G.E., Richard, P.E., Shaikh, T.H., Desselle, T.D., Hoppens, C.L., and Deininger, P.L. (1990) *Nucleic Acids Res.*, **18**, 6793−6798.
7. Ullu, E., Murphy, S., and Melli, M. (1982) *Cell*, **29**, 195−202.
8. Rogers, J. (1983) *Nature*, **301**, 460 .

9. Slagel, V., Flemington, E., Traina-Dorge, V., Bradshaw Jr., H., and Deininger, P.L. (1987) *Mol. Biol. Evol.*, **4**, 19–29.
10. Britten, R.J., Baron, W.F., Stout, D.B., and Davidson, E.H. (1988) *Proc. Natl. Acad. Sci. USA.*, **85**, 4770–4774.
11. Jurka, J., and Smith, T. (1988) *Proc. Natl. Acad. Sci. USA.*, **85**, 4775–4778.
12. Labuda, D., and Striker, G. (1989) *Nucleic Acids Res.*, **17**, 2477–2491.
13. Quentin, Y. (1988) *J. Mol. Evol.*, **27**, 194–202.
14. Willard, C., Nguyen, H.T., and Schmid, C.W. (1987) *J. Mol. Evol.*, **26**, 180–186.
15. Deininger, P.L., and Slagel, V.K. (1988) *Mol. Cell. Biol.*, **8**, 4566–4569.
16. Matera, A.G., Hellmann, U., and Schmid, C.W. (1990) *Mol. Cell. Biol.*, **10**, 5424–5432.
17. Batzer, M.A., and Deininger, P.L. (1991) *Genomics*, **9**, 481–487.
18. Matera, A.G., Hellmann, U., Hintz, M.F., and Schmid, C.W. (1990) *Nucleic Acids Res.*, **18**, 6019–6023.
19. Ausabel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1987) Current Protocols in Molecular Biology. John Wiley & Sons, N.Y., N.Y.
20. Friezner Degen, S.J., Rajput, B., and Reich, E. (1986) *J. Biol. Chem.*, **261**, 6972–6985.
21. SAS Institute Inc. (1963) SAS users guide edition v5.0. SAS Institute Inc., Cary, N.C.
22. Lipman, D.J., and Pearson, W.R. (1985) *Science*, **227**, 1435–1441.
23. Hutchison III, C.A., Hardies, S.C., Loeb, D.D., Shehee, W.R., and Edgell, MH. (1989) In Howe, M and Berg, D. (eds.) Mobile DNA. ASM Press, Washington D.C., pp. 593–618.
24. Flemington, E.H., Bradshaw, H., Traina-Dorge, V., Slagel, V., and Deininger, P. (1987) *Gene*, **52**, 267–277.
25. Lee, M.G.-S., Lommis, L., and Cowan, N.J. (1984) *Nucleic Acids Res.*, **12**, 5823–5836.
26. Stoppa-Lyonnet, D., Carter, P.E., Meo, T., and Tosi, M. (1990) *Proc. Natl. Acad. Sci. USA.*, **87**, 1551–1555.
27. Daniels, G.R., and Deininger, P.L. (1985) *Nucleic Acids Res.*, **13**, 8939–8954.
28. Kariya, Y., Kato, K., Hayashizaki, Y., Himeno, S., Tarui, S., and Matsubara, K. (1987) *Gene*, **53**, 1–10.
29. Cann, R.L., Stoneking, M. and Wilson, A.C. (1987) *Nature*, **325**, 31–36.
30. Miyamoto, M.M., Slightom, J.L., and Goodman, M. (1987) *Science*, **238**, 369–373.
31. Shen, M.R., Batzer, M.A., and Deininger, P.L. (1991) *J. Mol. Evol.*, Submitted.
32. Koop, B.F., Miyamoto, M.M., Embury, J.E., Goodman, M., Czelusniak, J., and Slightom, J.L. (1986) *J. Mol. Evol.*, **24**, 94–102.
33. Sawada, I., and Schmid, C.W. (1986) *J. Mol. Biol.*, **192**, 693–709.
34. Sawada, I., Willard, C., Shen, C.-K. J., Chapman, B., Wilson, A.C., and Schmid, C.W. (1985) *J. Mol. Evol.*, **22**, 316–322.
35. Botstein, D., White, R.L., Skolnick, M.H., and Davis, R.W. (1980) *Amer. J. Hum. Genet.*, **32**, 314–331.
36. Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C. Fujimoto, E., Hoff, M., Kumlin, E., and White, R. (1987) *Science*, **235**, 1616–1622.
37. Economou, E.P., Bergen, A.W., Warren, A.C., and Antonarakis, S.E. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 2951–2954.
38. Economou-Pachnis, A., and Tsichlis, P.N. (1985) *Nucleic Acids Res.*, **13**, 8379–8387.
39. Antonarakis, S.E., Oettgen, P., Chakravati, A., Holloran, S.L., Hudson, R.R., Feisee, L., and Karathanasis, S.K. (1988) *Hum. Genet.*, **80**, 265–273.
40. Mietus-Snyder, M., Charmley, P., Korf, B., Ladias, J.A.A., Gatti, R.A., and Karathanasis, S.K. (1990) *Genomics*, **7**, 633–637.