Vol. 171, No. 1, 1990 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 31, 1990 Pages 67-74

HUMAN LACTATE DEHYDROGENASE-B PROCESSED PSEUDOGENE: NUCLEOTIDE SEQUENCE ANALYSIS AND ASSIGNMENT TO THE X-CHROMOSOME

Kayoko Sudo[1,3], Masato Maekawa[1,4], Mary M. Luedemann[2], Larry L. Deaven[2] and Steven S.-L. Li[1,5]

[1]Laboratory of Genetics
National Institute of Environmental Health Sciences
National Institutes of Health
Research Triangle Park, North Carolina 27709

[2]Life Sciences Division
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

Received July 9, 1990

Two human genomic clones containing the lactate dehydrogenase-B processed pseudogene were isolated from two patients deficient in lactate dehydrogenase-B isozyme. The sequences of 3,287 nucleotides, including the pseudogenes and its flanking regions, from both clones were found to be identical except for three differences in the pseudogenes. The sequences of 1,286 nucleotides from these two pseudogenes exhibited 93% homology with the cDNA sequence of the lactate dehydrogenase-B functional gene, and the pseudogene contained 75/76 base substitutions, 11/12 single-base deletions, and 5 single-base insertions. This pseudogene was mapped to the x-chromosome by dot-blot analysis using a probe for the pseudogene or its 5' flanking sequence.

In humans, the lactate dehydrogenase A (muscle), B (heart) and C (testis) polypeptide chains are encoded by three different genes (1-3). The LDH-A and LDH-C genes are closely linked on chromosome 11, while the LDH-B gene is located on chromosome 12 (4-5). The genomic organization of human LDH-A, LDH-B and LDH-C genes has been elucidated and their protein-coding sequences are interrupted by six introns at homologous positions (6-8). During the characterization of genetic mutations resulting in deficiency of the human LDH-B isozyme (9-10), we have isolated and partially sequenced two genomic clones containing the LDH-B processed pseudogene.

[3]Present Address: Department of Laboratory Medicine
Jikei University School of Medicine
3-25-8 Nishi-Shinbashi
Minato-ku, Tokyo 105, Japan.

[4]Present Address: Department of Laboratory Medicine
Hamamatsu University School of Medicine
Handa-cho 3600, Hamamatsu City 431-31, Japan.

[5]To whom correspondence and reprint requests should be addressed.

0006-291X/90 $1.50

Here we report the nucleotide sequence of the LDH-B pseudogene and its flanking regions as well as the chromosomal assignment of this pseudogene.

Materials and Methods

Isolation and characterization of genomic clones:

The genomic DNAs were isolated from white blood cells of patients with a deficiency of the LDH-B subunit (9-10), and genomic DNA libraries were constructed from the MboI partially-restricted DNA fragments inserted at the BamHI site of the lambda GEM-11 (Promega). The full-length human LDH-B cDNA (11), excluding the poly(A) tail, were labelled with (α-^{32}P)dCTP (Amersham) with the use of random primers (Boehringer-Mannheim). This probe was used to isolate genomic clones containing LDH-B gene-related sequences by plaque hybridization as described previously (12). The plaque DNAs bound to the plaque-screen-hybridization membrane (DuPont) were hybridized at 65°C overnight, and the filters were washed as recommended by the supplier. Autoradiography of the filters was done with Kodak XAR-5 film and DuPont Cronex Lighting-plus intensifying screen at -70°C. The positive plaques were subsequently purified to homogeneity through repeated rounds of screening. The DNA purified from the positive genomic clones was analyzed by restriction endonuclease mapping and Southern blotting, as described previously (12). The isolated DNA fragments were further cleaved and subcloned into M13 mp18 or mp19 bacteriophages (13). The M13 bacteriophages exhibiting positive hybridization to the human LDH-B cDNA probe were isolated and the nucleotide sequences of the inserted DNA were determined by the dideoxy chain termination method with the sequencing protocol modified to use deoxyadenosine 5′-[α-^{35}S]thio)triphosphate (14).

Chromosomal assignment:

Human chromosomes were isolated and flow-sorted into individual spots on nitrocellulose discs as previously described (15). These spot-blots were hybridized with ^{32}P-labelled probe of the LDH-B pseudogene or its 5' flanking sequence. The conditions for hybridization and washing have been described previously (4).

Results and Discussion

Nucleotide sequence analysis:

Two genomic clones (B117 and B216) exhibiting the strongest hybridization signal to the human LDH-B cDNA probe were purified from different genomic libraries of patients with an LDH-B isozyme deficiency. Both genomic clones were found to possess inserts of approximately 15 Kb each, and their DNA fragments of 3.3 Kb containing the LDH-B gene-related sequence were isolated from SalI and EcoRI cleavage fragments. Both DNA fragments were further characterized by restriction endonuclease HindIII cleavage and Southern blot analyses, and both exhibited identical restriction patterns, presumably containing the same LDH-B gene-related sequences (Fig. 1). The SalI-EcoRI DNA fragments of 3.3 Kb and their subfragments cleaved by HindIII were subcloned into M13 bacteriophages, and their nucleotide sequences were determined. The sequences of 3,287 nucleotides from both genomic clones B117 and B216 were found to be identical except for three differences (Fig. 2).

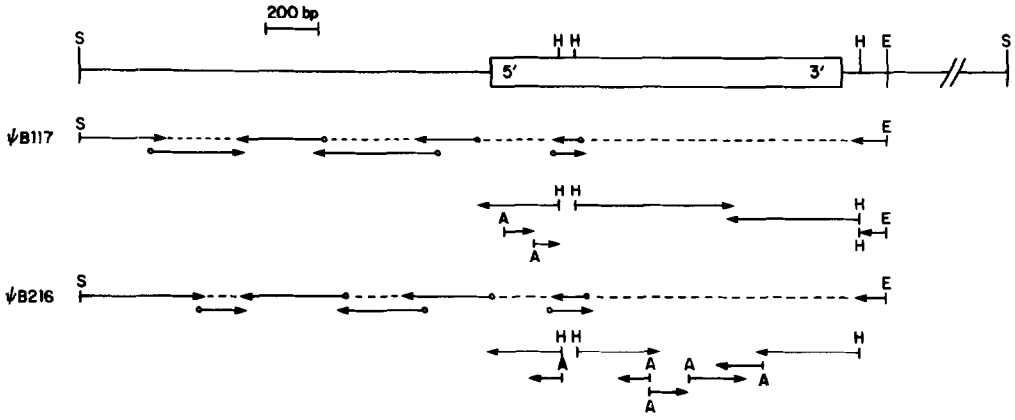

Fig. 1. Genomic structure of human LDH-B processed pseudogene along with restriction endonuclease map and nucleotide sequencing strategy.

The processed pseudogene is shown by the open block. The 5' and 3' flanking regions are indicated by solid lines. Genomic clones B117 and B216 were isolated from lambda GEM-11 libraries (Promega). The restriction endonuclease map was deduced from cleavage with EcoRI (R), HindIII (H) and SalI (S) (located at both ends of the genomic DNA inserts) and Southern blot analysis probed with the human LDH-B cDNA. The isolated SalI-EcoRI DNA fragments of 3.3 Kb were further cleaved with HindIII and subcloned into M13 mp18 or mp19 bacteriophages. The nucleotide sequences of the inserted DNAs were determined by the dideoxy chain termination method. The direction and length of sequencing are given. The open circle denotes the sequencing primer (17mer) synthesized according to the previously determined sequence.

Both genomic clones B117 and B216 contained the LDH-B gene-related sequences of 1,299 nucleotides, which are flanked by perfect direct repeats of 10 nucleotides. These sequences were found to have 93% homology with that of human LDH-B cDNA (Fig. 2A), and they possessed none of the introns which are present in the LDH-B functional gene (7). The sequence of 1,781 nucleotides upstream to the 5' direct repeat is very different from that of the putative promoter region of the human LDH-B functional gene. Therefore, the LDH-B processed pseudogene is transcriptionally inactive because of the absence of a promoter. Furthermore, both genomic clones B117 and B216 contained a substitution of G by A at the translation initiation codon ATG and an insertion of A right before the initiation site, as well as many other changes resulting in premature termination as discussed below. The 'coding' sequence of both genomic clones is incapable of producing a functional LDH-B protein. Therefore, the sequence of 1,299 nucleotides present in genomic clones B117 and B216 is an LDH-B processed pseudogene.

The processed pseudogene became nonfunctional at the time of its origin because of the absence of a promoter for transcription initiation. Thus, spontaneous mutations occurring in the pseudogene were apparently not subjected to functional (selective) constraints. It is of interest to know the rates and the types of mutations in the pseudogenes. Table I summarizes the nucleotide differences among 1,286 positions compared between both

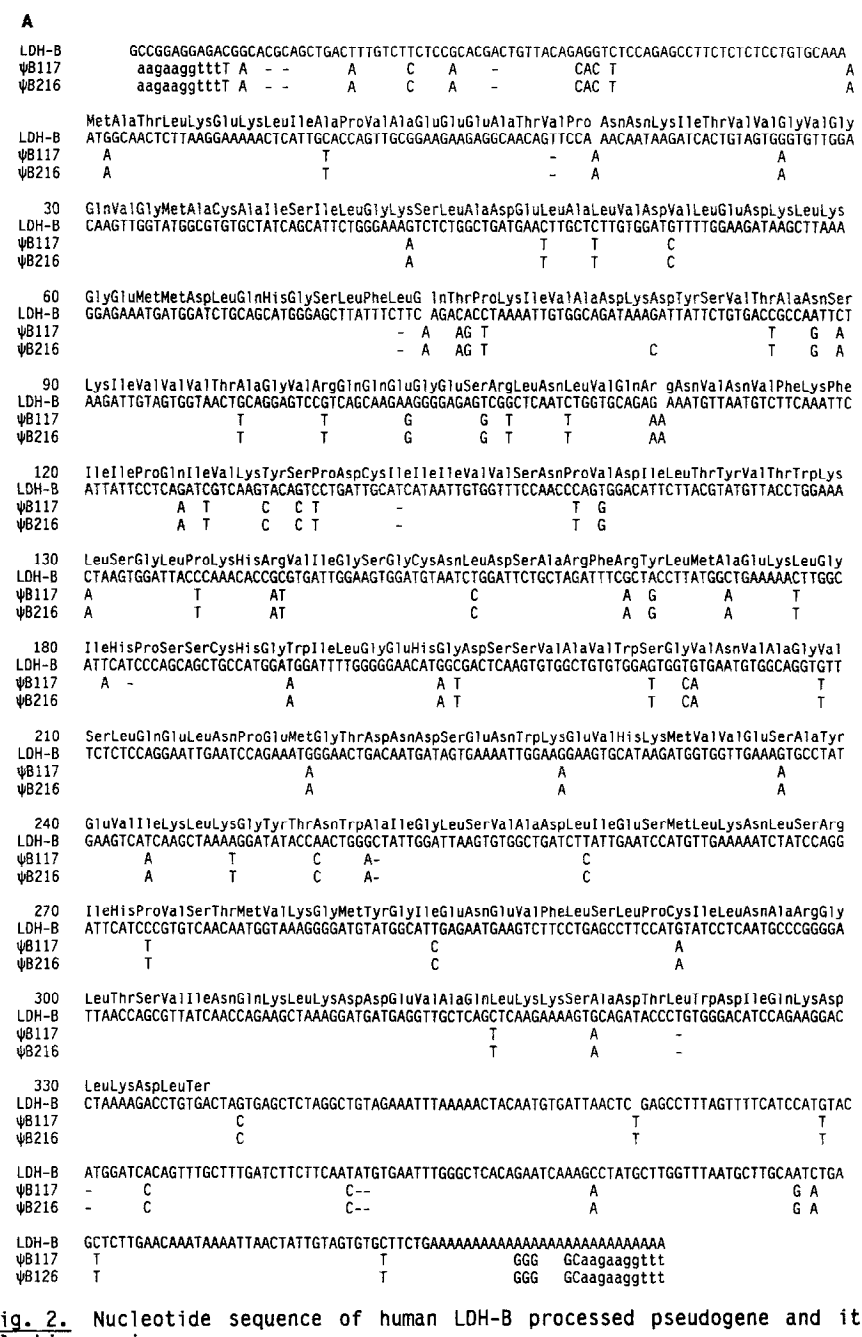

Fig. 2. Nucleotide sequence of human LDH-B processed pseudogene and its flanking regions.

A. Sequence comparison of human LDH-B processed pseudogenes present in genomic clones B117 and B216 with the cDNA sequence of the human LDH-B functional gene. The deduced amino acid sequence of the LDH-B cDNA is given above its nucleotide sequence. Different nucleotides between the pseudogene and cDNA sequences are given. Deletions are denoted by a hyphen, while insertions are shown by a gap in the cDNA sequence. Perfect direct repeats of 10 nucleotides are given in lower case.

B. The 5' flanking sequence of the LDH-B pseudogene.

C. The 3' flanking sequence of the LDH-B pseudogene.

LDH-B processed pseudogenes and the cDNA sequence, excluding poly (A) tail, of the LDH-B functional gene. There were 75/76 (82.4%) base substitutions, 11/12 (12.1%) single-base deletions and 5 (5.5%) single-base insertions.

B

```
      10        20        30        40        50        60        70        80        90
GTCGACTCTAGGCCTCACTGGCTAATACGACTCACTATAGGGAGCTCGATCCACTGGACTCAGCCTCCCAAAGTGCTGGGACTACAGGCATGATCCACCA  100
TGCCTGGCTCCATATAAACTTTAGAATTGTTTATTCTAAAGTTGTGAAAAATGTCATTAGTAAATTGATATTAACTGCATTGAATCCATAGATTGCTTTG  200
GGCAGTGTGGTCATTTTAATAATATTTATTCTTCTAATTGATGAGTATGGGACATTTTTCCATTTGTATCATCTACAATTTCCTTCATCAATGATTTCTA  300
GTTCTCCTTGTAGAGACTTTACCTTCTTACTTAAATGTATTCTGGCTTTCTTAAGGCCATTGTAAATAAAATTTCATTCTCAGTTTGGTTCTCAGCTTGA  400
TTGTTATTGGTGTATGAAAATGCTACTAATTTGTGTACACTGATTCTTTATTCTGAAACTTTGACTGCAGATTTTGTCAGTTCTAAGAGCCTTTTGGTGG  500
AGTCCTCAGAGTGTTCTAGGTATAAAATCATATCATCAGTAAAAATAGGTAATTTGATTTAAGTTCTTATTTGAATGCCTTTTATTGCTTTCACTTGTCT  600
CATTGCTCTGTAGGATTTCTTGAACTATGTTGAATAGGAGTGGTGAGAGTGCACAACCTCATCTTGTTTTAGTCTGAGAGAAAATGCTCCCCACTTTTTC  700
CATTTAGTATAATGTTGGCTGTGGGTTTGTAATATAAGGCTCTTATTATTTTGAGATACGTTTCTTTAATGTCTAGTTTGGTTTTTTAGTGTTTTTATCG  800
TGAAGGAATGTTGGACTTCATGGAATCCATTTTGTGCATCTATTGAGATGATTATATGGTTTTTATTTTTAAATCTGTTTATGTGGTGATTGATTTATTG  900
ATTTGTATGCCTTGAACCATCTTTGCATCTCTGGAGAAAATCTATTTGATCACAAGTATTATCTTTTGATGTGCTATTGGAATTGACTTGTTAGTATTTT 1000
GCTCAGGATTTTTTCATTTATGCTCACCTTGGATATTTGTCTGTAGTTTTCTTTCTTGTTATTGTGTGTTTGCCTGATTTGGTATCAGAGTGATACAGGC 1100
TTTGTAGTATGAGTTAGGGAGAAATCTCTCCTCTTTGGTATTTTTGAATAGTTTTAGTAGAATTGGTACTAGAGCCTGCATGTACATCCGGCAAAATTTG 1200
GCTATGAATCTGTTTAGTCCTGGTCTTCTTTGTTGAAAAATTTTACTACTGATTTGATTTAATTGCTCATTATTGGTCTGTTCAGGATTTCTGCTTCTAG 1300
TTTTAATTATCTATTTAAATTTTTGTGAGTACATAGTAGGTGTATATATTTGTGAAGTAAATTAGATGTTTTGGTACAGGCACATAATGCATAATAATCA 1400
CATTATAGAAAATGGAGCATCCATTCCCTGAAATATATATCTTTTGTGTTACAAACAATCCAATTATAACTCTTTTTAGTTATTCTAAAATGTGCAATTA 1500
AATTATTGTGTTAGACTGTTCTCACATGGTTATAAAGAAATACCTGAGACTGGGTAATTTATAAAGAAGGTTTAGTAAATTAGATGTTTTGGTACAGGCA 1600
CATAATGCATAATAATCACATTATAGAAAATGGAGCATCCATTCCCTGAAATATATATCTTTTGTGTTACAAACAATCCAATTATAACTCTTTTTAGTTA 1700
TTCTAAAATGTGCAATTAAATTATTGTGTTAGACTGTTCTCACATGGTTATAAAGAAATACCTGAGACTGGGTAATTTATA                    1781
```

C

```
      10        20        30        40        50        60        70        80        90
AATTGGCTCATGATTCTGCAGGTTGTACAGAAAGCATAGCATCATCTCTTCTGAGGAAGCTTCAGGAAGCTTCTCATCATGGCAGAAGGCAAAGGGAGAG  100
CCGTAACATCACATGGTGAAAGCAGGAGCAAGAGACTGAGACGAGAGGTGCTATACACTTTAAAATGACTAGCTCTCACAAGAATTC               187
```

Fig. 2-Continued

The 5' non-coding region appeared to have twice as many differences (17.6%) as those of the 'coding' (6.5%/6.7%) and 3' non-coding (6.2%) regions. The rate of mutations in the pseudogenes should be very close to the

Table I. Comparison of human LDH-B pseudogene and cDNA sequences

Types of changes	Numbers of changes in regions			
	5' NC[b]	'Coding'	3' NC	Overall
Base substitutions	9	57 (58[c])	9	75 (76)
Deletions[a]	3	5 (6)	3	11 (12)
Insertions	1	3	1	5
Total	13	65 (67)	13	91 (93)
Positions Compared	75	1005	209	1286
Differences (%)	17.6%	6.5% (6.7%)	6.2%	7.1% (7.2%)

[a] All deletions and insertions are single-base changes.

[b] NC means non-coding regions.

[c] LDH-B pseudogenes present in genomic clones B117 and B216 contained the same nucleotide sequence except that clone B117 had two more changes at 'codon' 180 and 181. The values in parentheses include the two additional changes in clone B117.

Table II. Frequencies of nucleotide substitutions in LDH-B pseudogenes[a]

	5' NC region				'coding' region				3' NC region				Overall			
	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G
A	-		1		-	1	1	4	-		1	1	-	1	3	5
T	1	-	1		2	-	5	2		-		1	3	-	7	2
C*[b]		1	-		4	11	-		2	2	-		6	14	-	
Cg		1	-			4	-				-			5	-	
*G	2		1	-	14	4	2	-		1	1	-	16	5	4	-
cG	1			-	2	1		-				-	3	1		-

a. The substitutions are from the nucleotides in the left column to the nucleotides in the row.

b. The C* indicates the C not followed by G, and the *G means the G not preceded by C.

intrinsic rate of spontaneous mutation because of the absence of functional constraints. Using the average substitution rate for mammalian globin pseudogenes, estimated to be 4.7×10^{-9} per nucleotide per year, if all of 91/93 nucleotide differences out of 1,286 positions compared (0.070762/ 0.072317 change per site) were assumed to be accumulated in the human LDH-B pseudogenes, then these pseudogenes present in genomic clones B216 and B117 should have existed in humans for 15.06 and 15.39 million years, respectively.

Table II indicates the relative frequencies of various types of base substitutions common in these two LDH-B pseudogenes. It has been proposed that the methylated CpG dinucleotide is a mutational 'hot spot', because about 90% of methylated C residues occur at the CpG dinucleotides and the

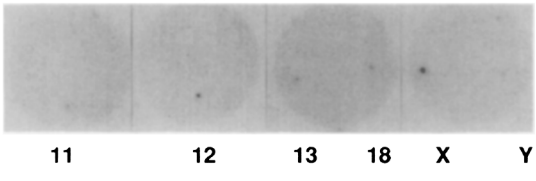

Fig. 3. Spot-blot analysis of flow-sorted human chromosomes using the LDH-B pseudogene probe.

The individual human chromosomes were flow-sorted directly on nitrocellulose discs. The separated chromosomal DNAs were denatured and hybridized with ^{32}P-labelled probe of the LDH-B pseudogene (HindIII DNA fragment of 1.2 Kb from genomic clone B216). The same spot-blots were also hybridized with the 5' flanking probe (SalI-HindIII DNA fragments of 1.8 Kb from genomic clone B117), and very similar results were obtained (data not presented).

methylated C, once deaminated, is not subject to the repair process (16). If this were the case in the pseudogene having no functional constraint, the CpG dinucleotides in a functional gene would be substituted mainly by TpG or CpA. In fact, a comparison between human LDH-B cDNA and the pseudogene sequences reveals that 11 out of 16 CpG dinucleotides contained mutations, namely, five changes to TpG, three to CpA, one to CpT, one deletion and one insertion. Excluding 9 changes at the CpG dinucleotides, 66 base substitutions present in the LDH-B processed pseudogene consisted of 42 (63.6%) transitions and 24 (36.4%) transversions. The frequencies of two transitions G to A and C to T are highest, and the two accounted for 45.5% of 63.6% transitions. The substitution pattern observed in the human LDH-B processed pseudogene is quite similar to those of human and mouse LDH-A processed pseudogenes as well as mammalian globin pseudogenes (17-19).

Chromosomal assignment:

In our previous study using the cDNA probes of human LDH-A, LDH-B and LDH-C (4), we have confirmed that human LDH-A and LDH-B genes are located on chromosomes 11 and 12, respectively. We also mapped the LDH-C gene and its related sequences on chromosome 11. Chromosomes 1, 2, 4, 9, and 10 were found to contain LDH-A gene-related sequences, whereas the X chromosome and chromosome 13 possess LDH-B gene-related sequences. In this investigation using the probes of the cloned LDH-B processed pseudogene and its 5' flanking sequence, the X chromosome exhibited the strongest hybridization signal, while chromosomes 12 and 13 gave much weaker signals (Fig. 3). These results clearly indicated that the cloned LDH-B processed pseudogene is located on the X chromosome. The cloned sequence of 3.3 Kb containing a human LDH-B processed gene will be of use as a molecular marker for the physical mapping and nucleotide sequencing of the entire human X chromosome.

Acknowledgments

We thank Drs. B. H. Judd, S. Ikawa, K. Machida and M. Kitamura for their support of this project; members of the Clinical Chemistry Laboratory at Toranomon Hospital and Ms. R. Jokyuu at Boshi Hokenin for their cooperation; Drs. F. Johnson and C. Teng for reading the manuscript; and Ms. N. Gore for preparing it. This investigation was supported in part by Japan Pathology Foundation for International Exchange.

References

1. Markert, C.L., Shakelee, J.B., and Whitt, G.S. (1975) Science 189, 102-114.
2. Li, S. S.-L., Fitch, W. M., Pan Y.-C. E. and Sharief, F. S. (1983) J. Biol. Chem. 258, 7029-7032.
3. Li, S.S.-L. (1989) Biochem. Soc. Transact. 17, 304-307.

4. Li, S.S.-L., Luedemann, M., Sharief, F.S., Takano, T. and Deaven, L.L. (1988) Cytogenet Cell Genet. 48, 16-18.
5. Kao, R.T., Wu, K.C., Law, M.L., Hartz, J.A., and Lau, Y.-F. (1988) Somatic Mol. Cell Genet. 14, 515-518.
6. Chung, F.Z., Tsujibo, H., Bhattacharyya, U., Sharief, F.S., and Li, S.S.-L. (1985) Biochem. J. 231, 537-541.
7. Takano, T., and Li, S.S.-L. (1989) Biochem. J. 257, 921-924
8. Takano, T., and Li, S.S.-L. (1989) Biochem. Biophys. Res. Commun. 159, 579-583.
9. Kitamura, M., Iijima, N., Hashimoto, F. and Hiratsuka, A. (1971) Clin. Chim. Acta. 34, 419-423.
10. Joukyuu, R., Mizuno, S., Amakawa, T., Tsukada, T., Nishina, T., and Kitamura, M. (1989) Clin. Chem. 35, 687-690.
11. Sakai, I., Sharief, F.S., Pan, Y.-C.E., and Li, S.S.-L. (1987) Biochem. J. 248, 933-936.
12. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning, a laboratory manual. Cold Spring Harbor Laboratory, New York.
13. Messing, J., Crea, R., and Seeberg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
14. Biggin, M.D., Gibson, T.J., and Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963-3965.
15. Deaven, L. L., van Dilla, M. A., Bartholdi, M. F., Carrano, A. V., Carm, L. S., Fuscoe, J. C., Gray, J. W., Hildebrand, C. E., Moyzis, R. K. and Pealman, J. (1986) Cold Spring Harb. Symp. Quant. Biol. 51, 159-167.
16. Lindahl, T. (1974) Proc. Natl. Acad. Sci. USA 71, 3649-3653.
17. Fukasawa, K. M. and Li, S. S.-L. (1987) Genetics 115, 177-184.
18. Tsujibo, H., Tiano, H.F., and Li, S.S.-L. (1985) Eur. J. Biochem. 147, 9-15.
19. Li, W.-H., Tanimura, M. and Sharp, P. M. (1987) J. Mol. Evol. 25, 330-342.