Novel actin crosslinker superfamily member identified by a two step degenerate PCR procedure

Timothy J. Byers[a,*], Alan H. Beggs[b], Elizabeth M. McNally[b], Louis M. Kunkel[b]

[a]*Department of Physiology and Biophysics, Indiana University School of Medicine, Indianapolis, IN 46202, USA*
[b]*Division of Genetics and Howard Hughes Medical Institute, The Children's Hospital, Department of Pediatrics, Harvard University School of Medicine, Boston, MA 02115, USA*

Received 19 April 1995; revised version received 16 June 1995

Abstract Actin-crosslinking proteins link F-actin into the bundles and networks that constitute the cytoskeleton. Dystrophin, β-spectrin, α-actinin, ABP-120, ABP-280, and fimbrin share homologous actin-binding domains and comprise an actin crosslinker superfamily. We have identified a novel member of this superfamily (ACF7) using a degenerate primer-mediated PCR strategy that was optimized to resolve less-abundant superfamily sequences. The ACF7 gene is on human chromosome 1 and hybridizes to high molecular weight bands on northern blots. Sequence comparisons argue that ACF7 does not fit into one of the existing families, but represents a new class within the superfamily.

Key words: Actin-binding protein; Multigene family; Polymerase chain reaction; Human brain; Dystrophin; β-Spectrin

1. Introduction

The actin cytoskeleton is central to cellular structure and motility; it provides stability or plasticity according to the needs of the cell. The organization of the actin cytoskeleton is regulated by a diverse group of proteins, including ones that crosslink actin filaments into bundles or networks. There are several families of actin-crosslinking proteins, and sequence analysis has revealed a homologous actin-binding domain in a large subset, referred to as the fimbrin, α-actinin, filamin family [1], or as an actin crosslinker superfamily ([2], also see [3]). This domain is found in members of the spectrin superfamily (β-spectrin, α-actinin, and dystrophin), and in ABP-280 (filamin), ABP-120 (gelation factor), as well as in fimbrin (plastin). Recent data suggest that the actin-binding domains of superfamily members are functionally homologous as well, and bind to the same sites on actin [3]. Despite homology in the actin-binding region, however, the proteins in these families serve distinct functions, range in size from 70 to 423 kDa and differ greatly in such basic features as their overall secondary structure [1,2,4].

Proteins of this actin crosslinker superfamily have proven to be important in hereditary diseases. Defects in dystrophin are responsible for the Duchenne and Becker muscular dystrophies [5], and defects in spectrin give rise to hereditary hemolytic anemias [6]. In addition, mutations in the α-actinin gene of *Drosophila* cause muscle weakness, paralysis, and atrophy [7]; ABP-280 expression can reverse motility defects in a human cell line lacking ABP-280 [8]; loss of ABP-120 has been shown to disrupt motility in *Dictyostelium* amoebae [9]; and mutations in the yeast fimbrin gene lead to defects in cytoskeletal organization and cellular morphogenesis [10].

We are seeking to identify additional members of this superfamily on the hypothesis that they are candidates for involvement in other hereditary disorders [11]. In previous studies, a protein with overall structural similarity to dystrophin, called utrophin or dystrophin-related protein, was identified by a DNA hybridization approach [12,13]. Past efforts in our laboratory using PCR amplification with degenerate primers resulted in the description of two new isoforms of human α-actinin [14]. However, the α-actinin transcripts were present in relatively high abundance, and we had difficulty identifying additional superfamily members that might have been present in lower amounts. Here we describe a novel member of this actin crosslinker superfamily identified by an enhanced degenerate primer-mediated PCR strategy.

2. Materials and methods

2.1. Primers and template DNA

Mixtures of oligonucleotides (degenerate primers) were designed from two conserved regions identified in sequence alignments of the actin-binding domains of actin crosslinker superfamily members. Because of considerable sequence divergence between superfamily members, it was necessary to limit the complexity in each mixture (i) by arbitrarily omitting some of the combinations necessary to encode all of the known superfamily members, and (ii) by using inosine (I) in some positions where maximal degeneracy was desired [15,16]. Each reaction employed a degenerate antisense primer referred to as 'WTIILR' (corresponding to dystrophin amino acid residues 542–559) that adds a *Not*I site to the product (underlined), [GCGGAGCGGCCGCI(C/T)(G/T)IA(A/G) (A/G/T) ATIA(A/G/T)I(G/T)TCCA]; and either a non-degenerate λgt10 primer near the polylinker, [GCCGTCGACTTGAGCAAGTTCAGCCTGGTT]; or a degenerate sense primer referred to as 'KTFTKW' (corresponding to dystrophin amino acid residues 261–280) that also adds a *Not*I site, [GGGGTGGCGGCCGCAGAA(A/G)AC(A/C/G/T)TT(C/T)AC(A/C/G/T) (A/G)(A/C)(A/G/C/T)TGG]. Lambda phage DNA was prepared from an oligo-dT-primed human fetal brain cDNA library in λgt10 [17] for use as template DNA in PCR amplification reactions. The library was amplified by plate lysate, phage were purified by PEG precipitation and cesium chloride gradient centrifugation, and DNA was isolated according to Sambrook, et al. [18].

2.2. Two-step PCR with degenerate primers

The PCR reactions (50 μl per reaction, 4 reactions) included 5 ng/μl template DNA (λgt10 library DNA containing cDNA inserts as above), primers (0.4 μM degenerate primer(s) or 0.2 μM λgt10 primer as indicated below), 10 mM Tris-HCl (pH 9), 50 mM KCl, 1.5 mM $MgCl_2$, and 0.001% gelatin. This mixture was incubated at 98°C for 5 min, then was brought down to 80°C for the addition of 0.2 mM each of dATP,

*Corresponding author. Fax: (1) (317) 274 3318.

SSDI 0014-5793(95)00722-9

dCTP, dGTP, dTTP, plus 0.025 units/μl Taq polymerase (Pharmacia). The mixture was then cycled (94°C, 30 s; 53°C, 4 min; 72°C, 4 min) 27 times. The first PCR step was carried out with the λgt10 and WTIILR primers. The products were precipitated with ethanol, resuspended, and separated on one lane of a 1.75% low-melting point agarose gel (NuSieve, FMC). Many bands were observed after the first amplification, and the region between 300 bp and 4,000 bp was sliced into 19 fractions (of roughly 50–200 μl per fraction) to separate major and minor bands. In the second PCR step, 0.5 μl of each melted gel slice was used as template DNA for 30 cycles of amplification in a 15 μl reaction volume with the KTFTKW and WTIILR primers (same conditions as above, except that the 72°C elongation step was for 1 min). This amplification was repeated after transferring 1 μl of the first amplification reaction into 50 μl of fresh reaction mix. The final reaction product was precipitated and electrophoresed on an agarose gel as above. Bands in the 300–400 bp range were excised from the gel, eluted using the Magic PCR Prep procedure (Promega Corp.), digested with *Not*I, and ligated into *Not*I-cut pBluescript SK II+ (Stratagene). White colonies were chosen by blue/white selection on X-Gal plates and DNA sequence was determined for subclones with inserts in the 300–400 bp range. Once the new actin-binding domain PCR product had been identified, it was labelled by PCR amplification in the presence of 10 μCi of [α-^{32}P]dCTP and was used for a hybridization screen of the same library used to prepare the template DNA for the cloning of additional cDNA's.

2.3. Sequence analysis

DNA sequence determination was carried out either manually using a Sequenase kit (United States Biochemical) or on an Applied Biosystems Inc. 373A sequenator. The DNA sequences of PCR products were used to conduct database homology searches at the National Center for Biotechnology Information using the BLAST network service [19]. DNA Sequences were analyzed using the Wisconsin Genetics Computer Group package [20]. The nucleotide sequence reported in this paper has been submitted to the GenBank/EMBL Data Bank with accession number L40626.

2.4. Northern blot analysis

^{32}P-labelled DNA probes were prepared by PCR amplification in the presence of 10 μCi of [α-^{32}P]dCTP in the place of unlabelled dCTP. Multi-tissue Northern blots containing 2 μg poly A$^+$ RNA from human tissues (Clontech) were hybridized according to the manufacturer's instructions.

2.5. Chromosome mapping

Southern blots containing *Hin*dIII digested DNA from a panel of nine human/hamster somatic cell hybrids [21] were screened by hybridization. In addition, selected monochromosomal human/rodent hybrid DNAs were amplified by PCR as described [22] and were separated on SSCP gels to distinguish the human PCR product from the homologous rodent locus [23].

3. Results

3.1. Strategy

The first amplification step utilized a primer corresponding to λgt10 vector sequence and an antisense actin-binding domain degenerate primer (WTIILR). The template was bacteriophage DNA purified from a λgt10 human brain cDNA library. The first step amplification products were gel fractionated in an attempt to achieve size resolution between different superfamily members by virtue of the distinctive 5′ end lengths of each family member. Some loss of size resolution was anticipated due to amplification of cDNA inserts that were not full length at their 5′ ends. This fractionation step allowed us to reduce the effects of competition by high abundance sequences such as β-spectrin and α-actinin. The second amplification step utilized a second actin-binding domain degenerate sense primer (KTFTKW) not used in the first round, along with the WTIILR antisense primer. This two-step procedure with different primer sets greatly increased the yield of superfamily cDNA clones as compared with two 30-cycle rounds of amplification containing the same two degenerate primers (not shown). In addition, because the spacing between these two primer sites is 300–400 bp in all known family members, the final PCR products were gel fractionated to select bands in this size range for further analysis.

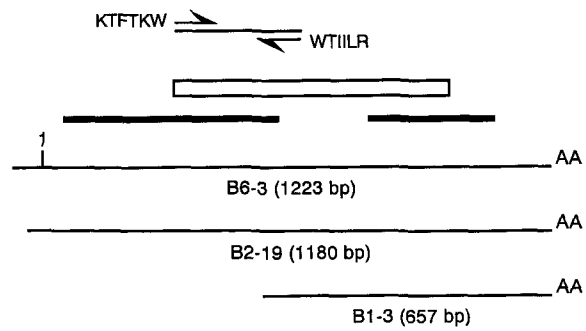

Fig. 1. Summary diagram of cDNA sequences. The ACF7 PCR product is shown on the top line with arrows representing the degenerate primers. The open box shows the region with homology to the actin crosslinker superfamily utilized for the sequence alignment in Fig. 2. The shaded bars show the extent of the probes used for northern blot hybridizations (bases 49–532 and 732–1026). The number 1 indicates the start of the region of exact match between library subclones B6–3 and B2–19, and the start of base pair numbering. The 'AA' indicates that each of these clones ended in a series of A residues (see text).

3.2. Results of the two-step degenerate primer-mediated PCR screen

We sequenced cDNA inserts of 61 randomly selected PCR subclones from 14 of the 19 fractions, and used the sequences to search the GenBank database. Of these, 22 matched β-spectrin, 8 were α-actinin cDNA's, and one cDNA encoded a novel protein with homology to the actin crosslinker superfamily (referred to as ACF7). Thirty PCR subclones did not show sequence similarity with the actin-binding domain of the actin crosslinker superfamily. Thus, 51% of the PCR subclones were specifically amplified from sequences encoding the expected actin-binding domains. Dystrophin cDNA clones were not recovered because they had been depleted in a previous screening and amplification of this library.

It was expected that a given transcript would tend to have the same length of 5′ sequence, and would cluster in a few fractions from the first gel fractionation. This was the general trend; β-spectrin cDNA's were found primarily in fractions originating from the top portion of the gel (10 of 11 PCR subclones from fractions 2 and 3 were β-spectrin). Similarly, all of the α-actinin cDNA's originated from slices in the bottom half of the gel, with 5 out of the 8 α-actinin subclones in fraction 19. The ACF7 cDNA was identified in fraction 5.

3.3. Analysis of library subclones

The ACF7 PCR product was labelled with ^{32}P and used as a probe to screen the original cDNA library. Six positive clones were isolated, and phage DNA was purified and digested with *Eco*RI. The digests were blotted to nitrocellulose paper and probed with the ACF7 PCR product. One *Eco*RI fragment was positive from each clone, and three positive-hybridizing frag-

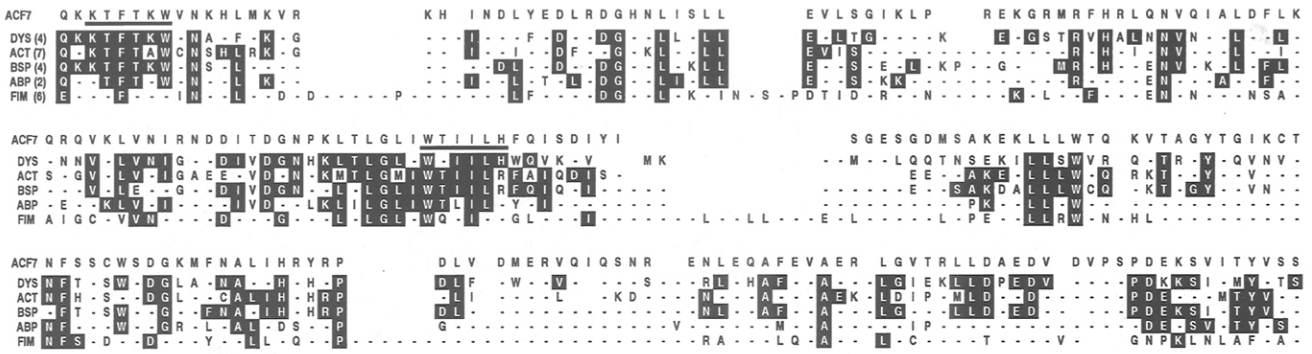

Fig. 2. Alignment of the ACF7 predicted amino acid sequence with consensus sequences from each actin crosslinker family. All of the peptide sequences listed below were aligned in the region of similarity with the PILEUP program. Gapped sequences written by the PILEUP program were used by the PRETTY program to generate consensus sequences of identical residues for each family (the number of family members used for each consensus appears in parentheses). Consensus residues that match the ACF7 sequence are shaded. Spaces indicate gaps introduced by the PILEUP program, and dashes indicate non-consensus residues within each family. Bars indicate the positions of the degenerate primers. Note that we have grouped ABP280 and ABP120 to form a consensus because they have structural similarities, but they are very different in size and may also be considered as two separate families. The sequences for each family were as follows (GenBank accession number and residue numbers in parentheses): 'ACF7', New actin-crosslinker: human (L40626; 80–292); 'DYS', dystrophin family: human (X14298; 17–234), mouse (M68859; 17–234), chicken (X13369; 21–238), and human utrophin (X69086; 33–249); 'ACT', α-actinin family: human ACTN1 (M95178; 33–244), ACTN2 (M86406; 40–251), ACTN3 (M86407; 47–258), Chicken (J03486; 34–245) and (X13874; 43–254), *D. discoideum* (Y00689; 24–237), and *D. melanogaster* (X51753; 36–247); 'BSP', β-spectrin family: human (J05500; 56–272), mouse (M74773; 56–272), *D. melanogaster* (M92288; 52–268), and (X53992; 26–246); 'ABP' family: human ABP280/filamin (X53416; 45–263), and *D. discoideum* ABP120/gelation factor (X15430; 14–221); 'FIM', fimbrin/plastin family: human I-plastin (L20826; 124–370), L-plastin (M22299; 65–312), and T-plastin (M22300; 65–312), rat T-plastin (X70706; 122–369), *S. cerevisiae* fimbrin (X63867; 141–384), and chicken fimbrin (X52562; 124–370).

ments of different sizes were subcloned into pBluescript SK II+ for further analysis. The largest positive-hybridizing *Eco*RI fragment, B6–3, was sequenced entirely on both strands. The entire length of B6–3 was open reading frame, ending in a stretch of A residues without a stop codon or polyadenylation signal, indicating the likelihood of aberrant priming at an internal series of A residues in the reverse transcription step of library construction. Clones B1–3 and B2–19 were sequenced at their ends. They both matched B6–3 exactly at their 3′ ends, also ending in a series of A residues (Fig. 1). At their 5′ ends, B2–19 and B6–3 began to match identically at base 67 of B6–3 and base 24 of B2–19. Because the short region of mismatch at the 5′ end of these clones may signal that one of them is a cloning artifact, we have designated the starting point of identical match as base 1 and we have submitted the sequence from 1–1157 to the GenBank/EMBL Data Bank. The 5′ end of B1–3 was identical with B6–3 beginning at base 501. For the purposes of this work, we have designated the first methionine (at base 52) in our open reading frame as residue 1 of the predicted amino acid sequence.

3.4. Sequence comparisons

To analyze relations between ACF7 and the homologous N-terminal regions of members of the actin crosslinker superfamily, we aligned the predicted amino acid sequence of ACF7 with consensus sequences of each family. The consensus sequences in Fig. 2 consist of residues that are identical in every available member of each family. As shown in Fig. 2, the ACF7 sequence matches many residues in each family consensus. There are also many positions that differ from each family. For each family consensus, there are positions that do not match ACF7, while another family does match at the same position, thus there is no one family whose consensus consistently agrees with ACF7. In addition, pairwise comparisons between ACF7 and the amino-terminal domains of the members of each family were consistently lower than comparisons between family members, even when highly divergent sequences from lower eukaryotes were included in the analysis (Table 1). These data provide evidence that the ACF7 product is a new class (and possibly a new family) within the actin crosslinker superfamily.

3.5. Analysis of expression

Northern blot hybridization revealed multiple high molecular weight bands after long exposure (data not shown). Though the bands were faint and not well resolved, they were reproducible in two independent blots. The bands were substantially larger than the 9 kb molecular weight marker and appeared to be larger in size than the dystrophin transcript (14 kb); we estimate them to be in the range of 14–20 kb. The highest signal strength was observed in lung, brain, liver and kidney, followed by pancreas, skeletal muscle, and heart, with the lowest level in placenta. Similar results were obtained on duplicate blots with each of the two probes diagrammed in Fig. 1.

Table 1
Pairwise sequence comparisons of ACF7 and actin-crosslinkers

Family	ACF7 vs. each family member	Minimum match within family
Dystrophin	49–51%	72%
α-Actinin	46–53%	64%
β-Spectrin	52–60%	75%
ABP120/ABP280	26–28%	41%
Fimbrin/plastin	39–42%	46%

The predicted amino acid sequence of ACF7 was compared with each actin crosslinker superfamily member (sequences as in Fig. 2) using the GAP program. The range of percentage identity for each family is listed in the first column. Pairwise comparisons were also carried out between the members of each family. The comparison that gave the minimum percent identity within each family appears in the second column.

Table 2
Segregation of ACF7[a] with human chromosomes in DNA from somatic cell hybrids

Chromosome	Hybridization pattern[b]				Discordant fraction[c]
	+/+	−/−	+/−	−/+	
1	1	8	0	0	0.00
2	0	5	0	2	0.28
3	1	4	0	3	0.38
4	0	5	1	2	0.38
5	1	5	0	3	0.33
6	1	4	0	3	0.38
7	0	4	1	4	0.56
8	0	3	1	5	0.67
9	0	6	1	2	0.33
10	0	4	0	3	0.43
11	1	5	0	2	0.25
12	1	4	0	3	0.38
13	1	5	0	3	0.33
14	1	3	0	4	0.50
15	1	7	0	1	0.11
16	1	3	0	5	0.56
17	0	5	1	3	0.44
18	0	4	1	4	0.56
19 and 19der[d]	1	1	0	7	0.78
20	1	1	0	7	0.78
21	1	4	0	4	0.44
22	1	5	0	3	0.33
X	1	6	0	1	0.13
Y	0	8	1	0	0.11

[a] The hybridization probe was ^{32}P-labelled B6-3 cDNA.
[b] Number of hybrids with indicated pattern of hybridization signal and chromosome. +/+ = hybridization signal and chromosome both present; −/− = hybridization signal and chromosome both absent; +/− = hybridization present but chromosome absent; −/+ = hybridization absent but chromosome present.
[c] Hybrids with a rearranged chromosome or in which the chromosome was present in fewer than 15% of cells were excluded for calculation of discordant fractions.
[d] This includes nine hybrids containing the der 19 translocation chromosome 19pter-q13::Xq24-qter and one hybrid containing 19qter-p13::Xq13-qter.

3.6. Human chromosomal locus

A human/hamster somatic cell hybrid mapping panel was screened by hybridization with B6–3. A single, predominant hybridizing *Hin*dIII fragment was present in control human DNA and in hybrid D4, but not in control hamster DNA (data not shown). Segregation analysis suggested that ACF7 is on chromosome 1 (Table 2). To confirm this, sequences corresponding to bases 78–280 of B6–3 were amplified by PCR from several independent mono- and dichromosomal hybrids [22]. The human sequence was present in a human/hamster hybrid that contained only human chromosomes 1 and X on a hamster background, and was not present in a hybrid that contained only the human × chromosome (data not shown). This confirms chromosome 1 as the location of the human ACF7 gene.

4. Discussion

Here we describe the actin-binding domain of a new actin crosslinker superfamily member. ACF7 maps as a single, novel locus on human chromosome 1. Sequence comparisons indicate that ACF7 does not fall into the existing family groupings, and probably represents a new class within this actin crosslinker superfamily. In the region of comparison, ACF7 bears the most similarity to β-spectrin, followed by α-actinin and dystrophin, with the least similarity to ABP120, ABP280, and fimbrin.

Expression of the ACF7 gene appears to give rise to a set of very large transcripts that show different tissue specificities, indicating the possibility of differential splicing and/or multiple promoters. Both of these mechanisms work in the dystrophin gene to give isoforms that differ dramatically in size [13,24]. It is apparent from the transcript size that considerable cloning remains to be done to isolate cDNA for, and to determine the complete ACF7 sequence.

Many variations of degenerate primer-mediated PCR have been used to isolate DNA sequences of related genes based on conserved regions of amino acid homology (see [25–27] for examples). Here we present a modified strategy to be tried in the event that a more straightforward approach fails to yield an adequate signal-to-noise ratio. The use of different primer sets in two PCR amplification steps increased the specificity and resulted in a high yield of superfamily members, while gel fractionation after the first amplification step accomplished a degree of discrimination between superfamily members, allowing the detection of ACF7, a lower abundance form.

Each family of this actin-crosslinker superfamily has a very distinctive and modular structure [1,2,4]. These distinctive features provide for variability in the length and rigidity of the crosslink, determining whether actin filaments are organized into tight bundles or loose gels. The only feature shared between all families is the actin-binding domain. It appears that these sequences have served as an evolutionary cassette to confer actin-binding activity to otherwise very different proteins. The product of the high molecular weight ACF7 transcripts may be a very large protein with a modular structure such as dystrophin, titin, or nebulin. Further analysis of the ACF7 gene and protein should give additional insight into its function as well as the evolutionary relations and structural diversity of this superfamily of actin binding and crosslinking proteins.

Acknowledgements: We thank Kiichi Arahata for isolating DNA from the λgt10 library. This work was supported in part by National Institutes of Health Grant NS 23740 to LMK. AHB was supported in part by the Charles H. Hood Foundation. LMK is an investigator of the Howard Hughes Medical Institute.

References

[1] Matsudaira, P. (1994) Semin. Cell Biol. 5, 165–174.
[2] Dubreuil, R.R. (1991) BioEssays 13, 219–226.
[3] Matsudaira, P. (1994) J. Cell Biol. 126, 285–287.
[4] Matsudaira, P. (1991) Trends Biochem. Sci. 16, 87–92.
[5] Monaco, A.P. and Kunkel, L.M. (1988) Adv. Hum. Genet. 17, 61–98.
[6] Palek, J. and Jarolim, P. (1993) Semin. Hematol. 30, 249–283.
[7] Roulier, E.M., Fyrberg, C. and Fyrberg, E. (1992) J. Cell Biol. 116, 911–922.
[8] Cunningham, C.C., Gorlin, J.B., Kwiatkowski, D.J., Hartwig, J.H., Janmey, P.A., Byers, H.R. and Stossel, T.P. (1992) Science 255, 325–327.
[9] Cox, D., Condeelis, J., Wessels, D., Soll, D., Kern, H. and Knecht, D.A. (1992) J. Cell Biol. 116, 943–955.
[10] Adams, A.E., Botstein, D. and Drubin, D.G. (1991) Nature 354, 404–408.
[11] Byers, T.J., Khurana, T.S., Beggs, A.H., Boyce, F.M., Lien, L.L. and Kunkel, L.M. (1991) in: Frontiers in Muscle Research (Ozawa, E., Masaki, T. and Nabeshima, Y., Eds.) pp. 341–350, Elsevier, Amsterdam.

[12] Love, D.R., Hill, D.F., Dickson, G., Spurr, N.K., Byth, B.C., Marsden, R.F., Walsh, F.S., Edwards, Y.H. and Davies, K.E. (1989) Nature 339, 55–58.
[13] Blake, D.J., Tinsley, J.M. and Davies, K.E. (1994) Trends Cell Biol. 4, 19–23.
[14] Beggs, A.H., Byers, T.J., Knoll, J.H.M., Boyce, F.M., Bruns, G.A.P. and Kunkel, L.M. (1992) J. Biol. Chem. 267, 9281–9288.
[15] Knoth, K., Roberds, S., Poteet, C. and Tamkun, M. (1988) Nucleic Acids. Res. 16, 10932.
[16] Erlich, H.A., Gelfand, D. and Sninsky, J.J. (1991) Science 252, 1643–1651.
[17] Feener, C.A., Koenig, M. and Kunkel, L.M. (1989) Nature 338, 509–511.
[18] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor.
[19] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[20] Devereux, J.R., Haeberli, P. and Smithies, O. (1984) Nucleic Acids. Res. 12, 387–395.
[21] Bruns, G., Stroh, H., Veldman, G.M., Latt, S.A. and Floros, J. (1987) Hum. Genet. 76, 58–62.
[22] Dubois, B.L. and Naylor, S.L. (1993) Genomics 16, 315–319.
[23] Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. (1989) Proc. Natl. Acad. Sci. USA 86, 2766–2770.
[24] Ahn, A.H. and Kunkel, L.M. (1993) Nature Genet. 3, 283–291.
[25] Grompe, M., Pieretti, M., Caskey, C.T. and Ballabio, A. (1992) Genomics 12, 755–760.
[26] D'Esposito, M., Pilia, G. and Schlessinger, D. (1994) Hum. Mol. Genet. 3, 735–740.
[27] Lopez, M., Oettgen, P., Akbarali, Y., Dendorfer, U. and Libermann, T.A. (1994) Mol. Cell. Biol. 14, 3292–3309.