Vol. 151, No. 1, 1988 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
February 29, 1988 Pages 487-492

A METHOD FOR SPECIFIC CLONING AND SEQUENCING OF HUMAN HPRT cDNA FOR MUTATION ANALYSIS.

D. Simpson, R.M. Crosby, and T.R. Skopek

Chemical Industry Institute of Toxicology
Research Triangle Park, NC 27709

Received January 29, 1988

A method has been developed for the specific amplification and cloning of human hprt cDNA which can be used for mutant sequence analysis. Messenger RNA is isolated from TK6 lymphoblasts and is used to produce a first strand cDNA with reverse transcriptase primed with oligo dT. Second strand synthesis and subsequent amplification of hprt sequences is accomplished using Thermus aquaticus DNA polymerase and hprt-specific primers in the polymerase chain reaction (PCR) procedure. Convenient restriction enzyme sites have been built into the 5' ends of the PCR primers to allow cloning of the hprt fragments in M13mp19. Dideoxy sequencing of hprt with specific primers can be carried out using either the PCR reaction product or fragments cloned in M13mp19 as substrate. This general cloning/sequencing method can be used to analyze hprt mutation in human cells obtained both in vitro and in vivo.

The advent of rapid DNA cloning and sequencing techniques has made it possible to analyze directly DNA base alterations induced by mutagenic agents. These alterations represent an unequivocal and permanent record of the final endpoint in the mutagenic pathway. Their analysis is extremely useful in understanding the mechanism of mutation, since it sets constraints on the possible DNA lesions involved in mutagenesis and on the enzymatic reactions acting on them.

Although a number of systems exist to study the sequence specificity of mutagenesis in E. coli (1,2,3,4) and recently in mammalian cells (5,6,7), none have been developed for the study of native sequences in human cells obtained in vitro or in vivo. The information generated by such a technique would be of interest both to molecular biologists interested in the mechanism of mutagenesis in human cells and to genetic toxicologists interested in the characterization of mutations in man.

We present here a general technique for the rapid sequence analysis of the hprt coding region in human cells. The technique uses the polymerase chain reaction (PCR) (8,9) to amplify hprt cDNA specifically for cloning and sequencing. The approach is demonstrated using human lymphoblasts grown in vitro, but should also be applicable for analysis of hprt sequences from human lymphocytes obtained in vivo (10,11,12,13). With minor modification, the technique can be used with any mammalian cell type grown in vitro.

0006-291X/88 $1.50

MATERIALS AND METHODS

Chemicals: The chemicals used in this study are as follows: guanidium isothiocynate, cesium chloride, isopropyl-β-D-thiogalacto-pyranoside (IPTG), 5-Bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal), restriction enzymes, M13mp19, T4 DNA ligase, T4 DNA kinase, M13 sequencing system, bacterial alkaline phosphatase (BAP) from Bethesda Research Laboratories; sodium citrate, β-mercaptoethanol, sodium pyrophosphate, N-lauryl-sarcosine from Sigma; oligo dT cellulose type 7 from Pharmacia Molecular Biology Division; cDNA synthesis system, Hybond-N from Amersham; Bacto-tryptone, Bacto-yeast extract, Bacto-Peptone, Bacto-Agar from Difco (American Scientific Products, MacGraw Park, IL); Polyethylene glycol from Aldrich. Chemicals used for transformation of bacteria were obtained from specific vendors as follows: hexamine cobalt chloride, potassium chloride, magnesium chloride-4H_2O gold label from Aldrich; calcium chloride-2H_2O from Fisher Scientific; MES from Sigma; dimethylsulfoxide from Mallinckrodt; dithiothreitol from Cal-Biochem. Oligonucleotides were synthesized on an Applied Biosystems 381 DNA Synthesizer. All reagents used to synthesize oligonucleotides were obtained from Applied Biosystems, Inc., Foster City, CA. 'Gene Amp' DNA amplification reagent kit was obtained from Perkin Elmer Cetus. A cDNA clone of human hprt was the generous gift of Dr. C.T. Caskey (Baylor College of Medicine, Houston, TX).

TK6 lymphoblast line: The cell line used was the TK6 human lymphoblastoid cell line previously described (14). Stock cells were maintained in RPMI 1640 supplemented with 15% heat-inactivated horse serum (both from GIBCO) in 75 cm^2 tissue culture flasks in a humidified incubator at 37°C with 6% CO_2. Cultures were diluted daily to 3 x 10^5 cells/ml.

Bacterial Strains: The following E. coli K12 strains were used: JM101 [supE, thi, Δ(lacproA,B)/F', traD36, proA,B, (r_k^+, m_k^+) lacI^q Z M15]; DH5α[end A1, hsdR17 (r_k^-, m_k^+) sup E44, thi-1, λ$^-$, rec A1, gyr A96, rel A1, θ 80d lac Z M15].

Media and Buffers. YT: 8 g/l Bacto-tryptone, 5 g/l bacto-yeast extract, 5 g/l NaCl; SOB: 20 g/l Bactotryptone, 5 g/l Bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$; YT plates: YT media plus 15 g/l Bacto-Agar.

RNA Isolation: RNA is isolated from the TK6 cell line using a guanidinium/cesium chloride extraction procedure (15). RNA is ethanol precipitated, rinsed with ethanol, air dried, and then resuspended in 400 μl H_2O. Approximately 100 μg total RNA is isolated from 100 ml of cell culture. Incubation of RNA preparations at 37°C without degradation of the RNA confirms that preparations are free of RNAse contamination.

mRNA Isolation: Oligo-dT chromatography is used to isolate biologically active mRNA from TK6 total RNA. A 1 ml tuberculin syringe column containing approximately 0.5 mg dry weight of oligo dT cellulose is prepared by washing with three volumes of 0.1 M KOH followed by five volumes of application buffer (500 mM KCl, 10 mM Tris-HCl pH 7.5). The RNA sample is resuspended in sterile H_2O and heated to 70°C for 5 min, placed on ice for 5 min, and brought to room temperature. A 1/5X volume of 5X application buffer is added to the sample just prior to loading on the column. The effluent from loading is collected and re-loaded to insure full retention of the mRNA. The column is then washed with two volumes of application buffer and two volumes 1st elution buffer (100 mM KCl, 10 mM Tris-HCl, pH 7.5). The mRNA is eluted from the column with two 0.5 ml volumes of 2nd elution buffer (10 mM Tris-HCl, pH 7.5). These fractions are then ethanol precipitated overnight at -20°C. Yields range from 2 to 8 μg mRNA per 100 μg total RNA.

cDNA First Strand Synthesis: First strand cDNA synthesis is performed according to the Amersham cDNA Synthesis System using oligo dT as a primer and reverse transcriptase. [^{32}P] dCTP is included in the reaction to monitor synthesis efficiency. After first strand synthesis each reaction is brought to 200 mM EDTA and treated at 60°C for 60 min in 0.2 N NaOH to digest the RNA. The mixes are neutralized by the addition of 3.0 M Na acetate (pH 5.0) to a final concentration of 0.3 M, ethanol precipitated, resuspended in water, and run over a Sephadex G-50 spin column (Boehringer Mannheim) that has been equilibrated with H_2O. Samples are then dried down overnight under vacuum.

Polymerase Chain Reaction (PCR) Amplification: Single-stranded hprt cDNA is specifically converted to double-stranded DNA and amplified in vitro using the polymerase chain reaction (PCR) technique using Thermus aquaticus (Taq) DNA polymerase under standard PCR conditions as specified by Perkin Elmer Cetus. In this technique, a set of two primers complementary to the 3' ends of the region to be amplified are added to the reaction mix in a

PCR Primers	Sequencing Primers
A $_{85}$CCGGtCGaCTCCGTTATG$_{103}$ (SalI)	1A $_{183}$ATAATGATTAGGTATGC$_{167}$
B $_{378}$TCCgtcGACTGTAGATTTTATCA$_{400}$ (SalI)	1B $_{414}$ATAGCTCTTCAGTCTGAT$_{397}$
C $_{429}$TGAgaatTCATTACAATAGCTCTTCA$_{404}$ (EcoRI)	1C $_{491}$AAGACATTCTTTCCAGTT$_{474}$
D $_{777}$AACTCAACTTGAAtTCTCATCTTA$_{754}$ (EcoRI)	1D $_{305}$TTGAGCACACAGAGGGCT$_{288}$
	1E $_{245}$AGACGTTCAGTCCTGTCC$_{228}$
	1F $_{654}$GTCTGGAATTTCAAATCC$_{637}$
	1G $_{745}$ATTTTGCTTTTCCAGTTTC$_{727}$

Base numbers are those assigned by Jolly et al. (18). All sequences are shown 5' to 3'. Lower case letters denote mismatched bases with hprt bases. Restriction enzyme sites are underlined.

Figure 1. Sequence and location of hprt primers used in PCR amplification and DNA sequencing.

large molar ratio (1.0 μg of each primer, 1.0 μg cDNA). The DNA is heat denatured at 94^{o}C (1 min) and allowed to cool at 37^{o}C (2 min). The primers anneal to the DNA and are then extended with Taq polymerase at 72^{o}C (3 min). This process is repeated several times, resulting in the exponential growth of the fragment of interest (8,9). A variety of primers have been synthesized for PCR amplification of hprt (Figure 1). These can be used in different combinations to amplify the first half (A and C), the second half (B and D), or the entire hprt coding region (A and D). The amplified samples are run over a Sephadex G-50 spin column equilibrated with H_2O to remove any impurities that might inhibit sequencing or ligation. Samples are dried down overnight under vacuum.

Cloning: The 5' ends of the primers have convenient restriction enzyme sites built into their sequence to provide easy cloning of the amplified sequence (Figure 1). These do not interfere with either primer annealing or extension during the PCR reaction. The samples are cut with EcoRI and Sal I (buffers as specified by vendor), extracted with phenol and ether, and ethanol precipitated. The cloning vector used is M13mp19, which has been previously cut with Sal I and EcoRI, treated with bacterial alkaline phosphatase, and phenol extracted. The ligations are carried out in 100 μl ligation buffer from BRL containing 0.1 μg vector and half of the PCR reaction. The samples are incubated overnight at approximately 14^{o}C with 1 unit of T4 DNA ligase per ligation. Before transformation the samples are phenol-extracted once, ether-extracted once, and ethanol precipitated. Each ligation is resuspended in 30 μl of 100 mM KCl, 45 mM $MnCl_2$, 10 mM $CaCl_2$, 3 mM $CoCl_3$, 10 mM K-MES.

The transformation procedure used is that of D. Hanahan (16) using E. coli strain DH5α. One third of each ligation was used per transformation with three transformations done for each ligation. Transformations are incubated 30 min on ice and heat shocked for 3 min at 37^{o}C. Two 100 μl aliquots of the transformations are plated on YT plates in three ml of top agar, with 200 μl of a log phase E. coli JM101 culture, 100 μl of a 2% X-gal solution, and 20 μl of 100 mM IPTG. The plates are incubated overnight at 37^{o}C.

White plaques (50 to 200) are purified and blotted onto circular nylon filters (Hybond N) in a grid pattern. Duplicate filters are probed with two different 15-mer oligonucleotides homologous to internal hprt sequences (unique from primers used for PCR). This is an important step since up to 90% of the clear plaques isolated contain only PCR primers 'forced' into the cloning site

because of their relatively high molar concentration during ligation. The hybridization protocol, labeling of the oligonucleotides, and the washing techniques are according to D. Woods (17). Plaques that are positive for both probes are then grown in JM101 and their DNA isolated for sequencing.

DNA Sequencing: We have synthesized a variety of primers located thoughout the hprt coding region to be used in dideoxy sequencing (Figure 1). The entire hprt coding region can be sequenced with these primers. Dideoxy sequencing is carried out as specified in the BRL M13 sequencing system with minor modifications. Dideoxy concentrations used were 0.125 mM (ddG and ddA), 0.25 mM (ddC), and 0.5 mM (ddT); 7-deaza-dGTP was used in place of dGTP. When sequencing the PCR product directly, the hprt fragment is electroeluted from a 1% agarose gel to remove PCR primers.

RESULTS

Thirty to forty cycles of PCR amplification of 1 μg of first strand TK6 cDNA results in bands visible on agarose gels stained with ethidium bromide. Southern blot (Figure 2) or sequence analysis reveals that the DNA bands contain hprt sequence. A single set of primers (A and D) can be used to amplify the entire sequence.

Due to the tremendous amplification power of the PCR reaction, great care must be taken to guard against accidental contamination of solutions and samples with exogenous hprt DNA (i.e. from hprt probes or previous PCR reactions). Negative controls are constantly run to monitor the purity of solutions used in the procedure (Figure 2). RNA samples treated with boiled RNAse and used in the procedure did not yield bands upon PCR amplification (data not shown).

We have also successfully cloned the hprt PCR product in M13mp19 for storage and further amplification. Cloning was accomplished using the EcoRI and Sal I restriction sites built into the PCR primers. Several hundred white plaques are obtained per cloning.

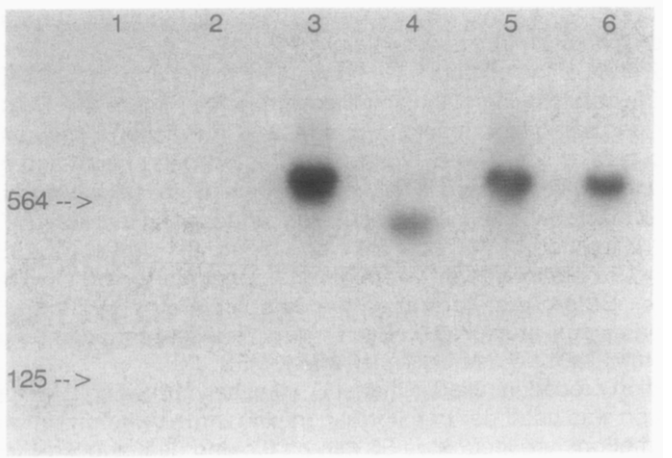

Figure 2. Southern blot analysis of PCR reactions products. All PCR reactions were carried out for forty cycles. Lane 1: primers A and D, no TK6 cDNA; Lane 2: primers B and D, no TK6 cDNA; Lane 3: primers A and D, with TK6 cDNA; Lane 4: primers B and D, with TK6 cDNA; Lane 5: primers A and D with TK6 cDNA plus 10^4 copies of hprt cDNA; Lane 6: primers A and D with 10^4 copies of hprt cDNA (no TK6 cDNA). Reactions were run on a 1.5% agarose gel, transferred to nylon membrane, and probed with ^{32}P-labelled hprt cDNA. Arrows indicate location of lambda-HindIII size markers.

The entire hprt coding region from TK6 cells has been determined and found to be identical to the hprt sequence published by Jolly et al. (18).

DISCUSSION

The analysis of mutations at the nucleotide level in mammalian cells requires a) a hemizygous target gene, b) a selection scheme to isolate cells carrying mutations in the target gene, and c) and method to retrieve and amplify the target gene DNA for sequence analysis. The hprt gene is a good candidate for this type of study. Since it is X-linked, hprt is hemizygous in males and effectively hemizygous in females due to the inactivation of one of the X chromosomes. Selection conditions (growth in 6-thioguanine) are available to isolate hprt$^-$ mutants both in vitro (19) and in vivo (20). The present study offers an approach of retrieving hprt coding sequence for sequence analysis.

The extremely large size of the hprt gene (41 kb long with 9 exons; ref. 21) renders repetitive analysis of genomic sequences impractical. The use of cDNA in the present approach circumvents the problem of size, since the large noncoding introns have been spliced out by the cells during mRNA processing. However, since cDNA is utilized as the source of sequence information, only point mutations in the 657 base coding region of hprt can be isolated and analyzed. Insertions and deletions involving large portions of the hprt genomic gene must be characterized by Southern blot analysis (22,23). Most point mutations in introns which interfere with accurate mRNA splicing will probably be missed as well. However, these can be identified using Northern blot analysis.

One caveat must be observed when working with the cloned PCR products. It must be remembered that mutations will be induced in vitro by DNA polymerase during the PCR reaction. Although the error rate of the Thermus aquaticus enzyme is not known, one can estimate based on other polymerases that the rate will probably be between 10^{-5} to 10^{-3} mutations per base. This means that many of the amplified strands will carry mutations at the end of the PCR reaction. One should therefore work with pooled M13 isolates ($>>10$) when sequencing; polymerase-induced mutations at a base in a given isolate will be 'diluted out' by the wild-type sequence in the other isolates. Since the biologically relevant mutation was present in all of the original cDNA copies, it will be present in all isolates, and therefore will be detected. This argument also applies when sequencing directly from the PCR reaction; only the original mutation will be present in the majority of molecules and will be the only one detected.

With present rapid sequencing methods and the advent of automatic sequencers, the task of sequencing the entire coding region to locate the mutation is not labor prohibitive. However, it would be prudent to use the RNAase A mismatch mapping technique (24) to locate at least some of the mutations to reduce the amount of sequencing necessary.

The technique presented should represent a general approach for the sequence analysis of mutation in human cells generated in vitro or in vivo. The use of other primers specific to hprt

from other species or to different genes is also possible and would expand the utility of the cDNA/PCR approach.

Note: A more detailed procedure and test samples of PCR primers can be obtained from the authors upon request.

REFERENCES

1. Coulondre, C., and Miller, J.H. (1977) J. Mol. Biol. 117, 577-594.
2. Fuchs, R.P.P., Schwartz, N., and Daune, M.P. (1981) Nature (London) 294, 657-659.
3. Skopek, T.R., and Hutchinson, F. (1982) J. Mol. Biol. 159, 19-33.
4. Richardson, K.K., Richardson, F.C., Crosby, R.M., Swenberg, J.A., and Skopek, T.R. (1987) Proc. Natl. Acad. Sci. 84, 344-348.
5. Calos, M.P., Lebkowski, J.S., and Botchan, M.R. (1983) Proc. Natl. Acad. Sci. 80, 3015-3019.
6. Ashman, C.R., and Davidson, R.L. (1987) Proc. Natl. Acad. Sci. 84, 3354-3358.
7. Nalbantoglu, J., Phear, G., and Meuth, M. (1987) Mol. Cell. Biol. 7, 1445-1449.
8. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) Science 230:1350-1354.
9. Scharf, S.J., Horn, G.T., and Erlich, H.A. (1986) Science 233, 1076-1078.
10. Bradley, W.E.C., Gareau, J.L.P., Seifert, A.M., and Messing, K. (1987) Mol. Cell. Biol. 7, 956-960.
11. Albertini, R.J., O'Neill, J.P., Nicklas, J.A., Heintz, N.H., and Kelleher, P.C. (1985) Nature 316, 369-371.
12. Skulimowski, A.W., Turner, D.R., Morley, A.A., Sanderson, B.J.S., and Haliandros, M. (1985) Mutat. Res. 162, 105-112.
13. Turner, D.R., Morley, A.A., Haliandros, M., Kutlaca, R., and Sanderson, B.J. (1985) Nature 315, 343-345.
14. Liber, H.L., and Thilly, W.G. (1982) Mutat. Res. 94, 467-485.
15. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, pp. 194-195. Cold Spring Harbor Laboratory, New York.
16. Hanahan, D. (1985) in DNA cloning: A Practical Approach (D.M. Glover, Ed.), Vol. I, pp. 109-135 , Academic Press, Washington DC).
17. Woods, D. (1984) in Focus (N. Sasavage, Ed.), Vol. 6, No. 3, pp. 1-2, Bethesda Research Laboratories, Gaithersburg, MD.
18. Jolly, D.J., Okayama, H., Berg, P., Esty, A.C., Filpula, D., Bohlen, P., Johnson, G.G., Sively, J.E., Hunkapillar, T., and Friedmann, T. (1983) Proc. Natl. Acad. Sci. 80, 477-481.
19. Thilly, W.G., DeLuca, J.G., Furth, E.E., Hoppe IV, H., Kaden, D.A., Krolewski, J.J., Liber, H.L., Skopek, T.R., Slapikoff, S., Tizard, R.J., and Penman, B.W. (1980) in Chemical Mutagens (F.J. de Serres and A. Hollaender, Eds.), Vol. 6, pp. 331-361, Plenum Press, New York.
20. Albertini, R.J., O'Neill, J.P., Nicklas, J.A., Heintz, N.H., and Kelleher, P.C. (1982) Proc. Natl. Acad. Sci. 79, 6617-6621.
21. Kim, S.H., Moores, J. C., David, D., Respess, J.G., Jolly, D.J., and Friedmann, T. (1986) Nuc. Acids Res. 14, 3103-3118.
22. Liber, H.L., Call, K.M., and Little, J.B. (1987) Mutat. Res. 178, 143-153.
23. Fuscoe, J.C., Fenwick Jr., R.G., Ledbetter, D.H., and Caskey, C.T. (1983) Mol. Cell. Biol. 3, 1086-1096.
24. Myers, R.M., Larin, Z., and Maniatis, T. (1985) Science, 230, 1242-1246.