ABSTRACT

We describe a method for direct cycle sequencing of PCR fragments amplified from genomic DNA or cDNA. DNA sequencing template is amplified using PCR and oligonucleotide primers flanking the region of interest. The amplified fragment is directly cycle sequenced using fluorescent sequencing primers, Sanger dideoxy sequencing chemistry and an enzyme mixture of a mutant Taq DNA polymerase and thermostable pyrophosphatase. The sequence ladders produced are analyzed on a real-time, automated four-color sequencing system. The method produces sequence ladders from unpurified PCR fragments of sufficiently high quality such that heterozygotes can be reproducibly detected and identified by software that recognizes signal-strength patterns indicative of mixed-base positions.

INTRODUCTION

Alterations in the wild-type sequence of a gene can have biological and clinical significance, and many methods have been developed for detecting mutations in genes. However, the only method that inherently provides complete information of the nature, location and predicted protein change of a genetic mutation is DNA sequencing (5). Recently, techniques for the direct sequencing of polymerase chain reaction (PCR) fragments have been developed. A common method utilizes biotin on one of the two PCR amplification primers and subsequent capture on streptavidin-coated magnetic beads (4,9). The captured fragment is then purified, denatured and sequenced with a modified form of T7 DNA polymerase such as Sequenase®. While this method gives high-quality sequence data from which heterozygotes can be readily identified, it is labor intensive, expensive and requires large amounts of sequencing template.

Direct PCR sequencing methods using thermostable DNA polymerases have also been described (6,11). Sequencing using thermophilic DNA polymerases has the advantage of needing no chemical denaturation, since cycle sequencing can be done with heat denaturation. Cycle sequencing methods using thermostable DNA polymerases also require less template, since termination reactions are linearly amplified in the cycling process. Two methods of cycle sequencing are routinely used: labeled sequencing primers with unlabeled dideoxynucleotides and unlabeled sequencing primers with labeled dideoxynucleotides. However, both these cycle sequencing methods have had the disadvantage of producing uneven termination patterns. This makes determination of heterozygotes difficult, since the dideoxynucleotide incorporation of the two bases at a heterozygous position is frequently unequal. Also, some thermophilic DNA polymerases have a 5' to 3' exonuclease activity. Substrates for this exonuclease activity are known to form during cycle sequencing, resulting in cleavage of the DNA template as well as the newly synthesized terminated strand (7,12,13). These cleavages show up as noise and false termination peaks (or stops) in the DNA sequence data.

Sequencing by enzymatic DNA polymerization is usually considered an irreversible reaction. However, all chemical reactions are governed by equilibrium between the forward and reverse reactions. The reverse reaction of DNA polymerization (called pyrophosphorolysis) occurs to a limited extent in Sanger dideoxy sequencing reactions. Sequence ladders in which pyrophosphorolysis occurs will often have regions in the ladders where termination peaks or bands have less signal strength or band intensity. In order to prevent termination peak dropouts due to pyrophosphorolysis, sequencing enzyme mixtures are often formulated with pyrophosphatase (14). The addition of pyrophosphatase catalyzes the removal of pyrophosphate, and the equilibrium of the extension reaction is thus shifted more favorably to the forward DNA polymerization reaction.

We have developed a simplified method for direct cycle sequencing of PCR fragments amplified from genomic DNA or cDNA. The use of a mutant
Table 1. PCR Primers Used for Amplification of hMSH2 Gene

<table>
<thead>
<tr>
<th>Exon 5 Primers</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EX5-5’</td>
<td>TGTAACGACGCCAGTTGCACTGGCTATAGGAAATCT</td>
</tr>
<tr>
<td>EX5-3’</td>
<td>CAGGAACAGCTATGACACCAACATTCTACATTTTAACCC</td>
</tr>
<tr>
<td>Exon 8 Primers</td>
<td>Sequence (5’ to 3’)</td>
</tr>
<tr>
<td>EX8-5’</td>
<td>TGTAACGACGCCAGTTGCACTGGCTATAGGAAATCT</td>
</tr>
<tr>
<td>EX8-3’</td>
<td>CAGGAACAGCTATGACACCAACATTCTACATTTTAACCC</td>
</tr>
<tr>
<td>Exon 12 Primers</td>
<td>Sequence (5’ to 3’)</td>
</tr>
<tr>
<td>EX12-5’</td>
<td>TGTAACGACGCCAGTTGCACTGGCTATAGGAAATCT</td>
</tr>
<tr>
<td>EX12-3’</td>
<td>CAGGAACAGCTATGACACCAACATTCTACATTTTAACCC</td>
</tr>
</tbody>
</table>

Taq enzyme (AmpliTaq® DNA Polymerase, FS) as a sequencing enzyme makes it possible to sequence heteroduplex PCR fragments without purification or chemical strand denaturation. Since the mutant AmpliTaq FS enzyme is more efficient in incorporating deoxyribonucleotides than wild-type AmpliTaq DNA Polymerase, deoxyribonucleotides carried over from the PCR into the sequencing reactions have little effect on sequencing results. The AmpliTaq FS enzyme gives more even termination patterns, fewer stops and less noise than cycle sequencing using wild-type Taq DNA polymerase (7,12,13). Also, the addition of a thermostable pyrophosphatase to the sequencing enzyme mixture eliminates termination peak dropouts due to pyrophosphorylation. The use of automated cycle sequencing with AmpliTaq FS enzyme and fluorescently labeled primers, in combination with appropriate software for data analysis, provides a rapid, non-radioactive mutation detection method.

We report on an investigation of the properties of this new approach. In this study, the hMSH2 gene was used as a model system for mutation analysis. The hMSH2 gene codes for a DNA mismatch repair protein, and inherited defects in the gene have been linked with the development of hereditary non-polyposis colorectal cancer (HNPPCC) (3,8). This gene is located on chromosome 2p and is comprised of sixteen exons. Multiple mutations responsible for HNPPCC have been identified and are widely interspersed throughout the gene (10). Our direct PCR sequencing and software analysis strategy was evaluated by applying the method to the analysis of mutations in three exons of the hMSH2 gene.

MATERIALS AND METHODS

PCR Amplification of hMSH2

Genomic DNA samples previously characterized as having mutations in hMSH2 were obtained from the Johns

Figure 1. Exon 5 point mutation. Shown is an exon 5 electropherogram of sample C76 in the sense direction. An A to T mutation is evident and is highlighted.
Hopkins Oncology Center, Johns Hopkins University, School of Medicine. The DNA samples were amplified using intronic PCR primers that were designed to flank exon 5, exon 8 and exon 12 of hMSH2. In order to facilitate sequencing, the primers were synthesized with the -21M13 sequence (tgtaaaacgacgccaggt) on the 5' end of the 5' amplification primer and the M13 reverse sequence (cag.gaa.aca.gct.atg.acc) on the 5' end of the 3' amplification primer. Amplifying with these priming sites on the ends of the PCR primers allowed all amplification products to be sequenced with the same fluorescently labeled -21M13 and M13 reverse sequencing primers. Table 1 shows the amplification primers used. Exons 5, 8 and 12 were amplified in a 50-µL PCR volume using 160 nM of each of the respective intronic PCR primers, 50 ng of genomic DNA, 100 µM of each dNTP, 2.0 U Taq DNA Polymerase (AmpliTaq®; Perkin-Elmer, Norwalk, CT, USA), 10 mM pH 8.8 Tris-HCl and 75 mM KCl. MgCl₂ concentrations were 1.5 mM for the exon 5 and exon 12 PCR and 3.5 mM for the exon 8 PCR. A hot start was done for all amplifications using AmpliTaq™ PCR Gem 50 wax beads (Perkin-Elmer). The lower reagent layer was in a volume of 25 µL and contained dNTPs, primers, dH₂O and an AmpliTaq™ PCR Gem 50. The wax bead was melted and solidified over the lower reagent layer by heating at 90°C for 1 min, followed by cooling to 4°C for 2 min. The upper reagent layer was a total volume of 25 µL and contained the buffer, Taq DNA polymerase and dH₂O. Optimization of the PCR buffer conditions was performed using an Opti-Prime™ PCR optimization kit (Stratagene, La Jolla, CA, USA). Thermal cycling was done in a GeneAmp® 9600 cycler (Perkin-Elmer). An

Table 2. Sequencing Primers Used to Sequence PCR Fragments of hMSH2

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5' to 3')</th>
<th>Region Sequenced</th>
</tr>
</thead>
<tbody>
<tr>
<td>-21M13</td>
<td>TGTAACGACGAGCAGGT</td>
<td>Sense direction of all PCR fragments</td>
</tr>
<tr>
<td>M13RP1</td>
<td>CAGGAAACAGCTATGACC</td>
<td>Antisense direction of all PCR fragments</td>
</tr>
</tbody>
</table>

Figure 2. Exon 8 point mutation. Shown is an alignment of sample K29 sense and antisense electropherograms. The antisense sequence (K29-R1) was reverse complemented before alignment. The T to G mutation is evident and is indicated by the arrows.
initial denaturation of 98°C for 20 s was followed by eight cycles of 98°C for 10 s, 60°C for 1 min and 70°C for 1 min. This was followed by thirty-two cycles of 96°C for 10 s and 68°C for 2 min. Initial cycling used a higher denaturing temperature to increase denaturing efficiency of the genomic DNA. A lower annealing temperature for initial cycles was done to increase priming and amplification efficiency. The later two-step cycling was done to increase amplification specificity. Amplification quality was evaluated by examining 10 μL of the amplified products on 1% agarose gels stained with 0.8 μg/mL ethidium bromide. Amplification yields were determined by spectrophotometry using an LS50B Luminescence Spectrometer (Perkin-Elmer) and PicoGreen™ ds DNA quantitation reagent (Molecular Probes, Eugene, OR, USA).

Yields of the amplified PCR products ranged from 5 to 10 ng/μL.

DNA Sequencing of Amplified PCR Products

Following amplification, 1 μL of the PCR product was diluted to 10 μL with dH₂O. The diluted PCR product was then directly sequenced with no purification using -21M13 and M13 reverse sequencing primers labeled with JOE, FAM, TAMRA and ROX. For each PCR product, four sequencing reactions were performed, one for each dideoxy termination reaction. Sequencing was done using the AmpliTaq DNA Polymerase, FS Dye Primer Core Kit (Perkin-Elmer Applied Biosystems Division [PE/ABI], Foster City, CA, USA). Four mixtures containing 80 nM dye-labeled primer, d/ddNTPs, AmpliTaq DNA Polymerase, FS containing thermostable pyrophosphatase, 80 mM Tris-HCl pH 9.0 and 2 mM MgCl₂ were made. The A sequencing reactions used JOE labeled primer; C sequencing reactions used FAM-labeled primer; G sequencing reactions used TAMRA-labeled primer; and T sequencing reactions used ROX-labeled primer. For the A and C reactions, 1 μL of diluted PCR product was added to 4 μL of the respective sequencing mixture. For the G and T reactions, 2 μL of diluted PCR product were added to 8 μL of the respective sequencing mixture. Following addition of PCR template to the reaction mixture, the samples were cycle-sequenced in a GeneAmp 9600 for fifteen cycles of 96°C for 20 s, 60°C for 30 s and 70°C for 30 s. This was followed by a further fifteen cycles of 96°C for 20 s and 70°C for 30 s.

<table>
<thead>
<tr>
<th>MSH2 exon 12</th>
<th>120</th>
<th>130</th>
<th>140</th>
<th>150</th>
<th>160</th>
</tr>
</thead>
<tbody>
<tr>
<td>J27-F1</td>
<td>ATGGAGCACC TGTCCCATAT GTACGACCAG CCACTTTTGG GAAAGGACAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J27-R1</td>
<td>ATGGAGCACC TGTCCCATAT GTACGACXAG CCACTTTTGG GAAAGGACAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>210</td>
</tr>
<tr>
<td>MSH2 exon 12</td>
<td>GAAGAATTAT TATTAAGAC ATCCAGACCAT GCTTGCTTGG AAGTCCAAGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J27-F1</td>
<td>GAAGAATTAT TATTAAGAC ATCCAGACCAT GCTTGCTTGG AAGTCCAAGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J27-R1</td>
<td>GAAGAATTAT TATTAAGAC ATCCAGACCAT GCTTGCTTGG AAGTCCAAGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Sense and antisense electropherograms of sample J27. Alignment of sample J27 in the sense and antisense directions. The antisense sequence (J27R1-M2) was reverse complemented before alignment. Arrows point to the C to T mutation.
After thermal cycling, the four sequencing reactions were pooled into 125 μL of 100% ethanol and centrifuged for 15 min. The ethanol was decanted and the reactions dried in a SpeedVac® (Savant, Holbrook, NY, USA) for 5 min. The reactions were resuspended in 8 μL of formamide/50 mM EDTA, pH 8.0 (5:1), denatured for 2 min at 95°C, and then 2 μL were loaded onto a 4% polyacrylamide gel (19:1 bis) containing 6 M urea and 1x TBE buffer. Electrophoresis was done for 7 hours at 1.6 kV, and electropherogram data were collected on an ABI 377 automated sequencer (PE/ABI).

**Data Analysis**

Sequence chromatogram files were analyzed using Factura™ software (PE/ABI). The software recognizes heterozygotes by identifying base positions that exhibit a secondary base signal strength greater than a selected percentage threshold of the primary base. Base positions that show a secondary base signal strength that is greater than the selected threshold are identified and flagged by the software. The signal strength ratio threshold was set at 50% for determination of all heterozygotes. After processing, chromatogram files were imported into Sequence Navigator™ software (PE/ABI) and aligned to wild-type reference sequence files. The marked base positions of potential mutations were visually examined to confirm heterozygote calls.

**RESULTS AND DISCUSSION**

The method described is a relatively simple means of generating and analyzing sequencing ladders that may contain heterozygote positions. It should be noted that optimization of the PCR amplifications was critical to obtain good results. Optimization of the PCR was facilitated by the use of a commercially available kit with a variety of buffer and salt conditions. It was also important to reduce the level of deoxynucleotides and amplification primers in the PCR to avoid introducing unwanted mutations.

![Figure 4. Sense and antisense electropherograms of sample JV. Alignment of sample JV in the sense and antisense directions. Reverse complementation of the antisense sequence JVR1 was done before alignment. Arrows point to the C to T heterozygous mutation.](image-url)
effects from amplification primers carried into the sequencing reactions.

The forward sequencing electro-

phorogram of sample C76 is shown in

Figure 1. Sample C76 is a 287-bp PCR

fragment that includes the entire coding

region of exon 5 of hMSH2 and intron

regions immediately surrounding it.
The red trace is the T termination data

and the green trace is the A termination
data. An A to T mutation is evident and

is highlighted. The text sequence above

electrophorograms shows the Inter-
national Union of Biochemistry (IUB)

code W for an A/T heterozygote. The

ratio of the A signal strength to the T

signal strength is 0.84, indicating that

the mutant T termination signal

strength is greater than the wild type A

signal strength. The mutation is at the

splice donor site of the 3' end of exon 5

of hMSH2. This mutation results in the
deletion of exon 5 from the translated

product (10).

Sample K29 is a 234-bp genomic

PCR fragment covering the coding re-

gion of exon 8 of hMSH2 and the in-

tron regions immediately flanking it.

Figure 2 shows sequencing electro-

phorograms of K29 in both directions.
The reverse sequence was reverse com-

plemented for ease of presentation. The

red trace is the T termination data and

the black trace is the G termination
data. The T to G mutation is highlight-
ed. The IUB code K for a G/T heter-

ozygote was determined by the soft-
ware a is shown in the text sequence

above the electrophogram. The ratio

of the signal strength of the forward

mutant G to wild-type T signal strength

is 0.80. Before reverse complementa-
tion, the ratio of the signal strength of

the reverse wild-type A to mutant C

signal strength is 0.66. The TGA muta-
tion at cDNA nucleotide 1441 creates a
termination codon. The mutation re-

sults in the truncation of the translated

product at codon 458 (10).

Samples J27 and JV are 373-bp ge-
nomic PCR fragments of exon 12 of

hMSH2 and the intron regions sur-

rounding it. Figure 3 shows the forward

and reverse sequence electrophero-

grams of J27. The reverse electrophero-

gram was reverse complemented for

alignment. The red trace is the T ter-

mination data, and the blue trace is the C
termination data. In sample J27, a C to

T mutation is highlighted. The IUB
code Y for a C/T heterozygote was
determined by the software and is

shown in the text sequence above the
electrophogram. The ratio of the for-
ward signal strength of the mutant T to

the wild type C is 0.77. The ratio of the

reverse signal strength of the mutant A
to the wild-type G is 0.83. The muta-
tion causes a proline to leucine amino

acid change in the translated product at
codon 622 (10).

Sample JV, shown in Figure 4, has a

C to T mutation and is highlighted in

both the forward and reverse chro-
matograms. The IUB code Y for a C/T

heterozygote is shown above the elec-
trophogram. The ratio of the forward

signal strength of the mutant T to wild-
type C is 0.59. The ratio of the reverse

signal strength of the mutant A to the

wild-type G is 0.94. This mutation cre-

ates a new splice donor site and results

in the deletion of codons 638 to 669 of

the translated product (10).

Figure 5 demonstrates the repro-
ducibility of the sequence ladder pat-
terns obtained. An alignment is shown

of the J27 and JV forward electrophero-
grams. Arrows point to the heterozy-
gous mutations. These positions are the

only positions in the electrophorograms

that differ by more than 20% of the

wild-type sequence.

CONCLUSIONS

Cycle sequencing using AmpliTaq

DNA Polymerase, FS allows direct se-

quencing of PCR products without the

need for purification or chemical dena-

turation. Using fluorescently labeled

primers in conjunction with AmpliTaq

FS enzyme and a thermostable py-

rophosphatase to sequence PCR frag-

ments produces data of sufficiently

high quality that heterozygotes can be

readily detected and identified using

relatively simple software tools. Auto-
mated fluorescent sequencing provides

data in a format that allows the ready

analysis of sequences for heterozygosi-
ty and pattern recognition. The identifi-
cation of mutations and the determina-
tion of their location are facilitated by

flagging of heterozygotes in electropho-

gram files and by alignment to wild-type

sequences. For increased

sensitivity and accuracy in heterozy-
gote detection, PCR amplification should be optimized. With the advent of long PCR and multiplex PCR, it should be possible to reduce the number of amplification reactions needed for determination of mutations in large genes such as hMSH2 and BRCA1 (1,2). The demonstration of successful detection of heterozygote positions in three different exons of hMSH2 indicate the potential generality of the approach. In fact, similar results have been obtained by applying the same methodology to the analysis of other human genes, including several HLA genes (data not shown). Further investigations are in progress.

ACKNOWLEDGMENTS

The authors thank Dr. Kenneth W. Kinzler and Dr. Bert Vogelstein of the Johns Hopkins Oncology Center, Johns Hopkins University School of Medicine, for providing the genomic DNA samples used for amplification. We thank Dr. Lisa Kalman, Dr. Alice Wang and Dr. David H. Gelfand in the Program in Core Research, Roche Molecular Systems, Inc., for providing research samples of AmpliTaq DNA Polymerase, FS and purified recombinant thermostable pyrophosphatase. We also thank Cheryl Heiner, Lin Zuo and Ellison Chen of the Advanced Center for Genetic Technology, Applied Biosystems Division of Perkin-Elmer, for insightful ideas and advice.

REFERENCES


Received 10 March 1995; accepted 29 September 1995.

Address correspondence to:
Robert B. Chadwick
Perkin-Elmer Corporation
Applied Biosystems Division
850 Lincoln Centre Drive
Foster City, CA 94404, USA