Ligase-based detection of mononucleotide repeat sequences

Monib Zirvi, Takamori Nakayama[1], Gregg Newman[2], Timothy McCaffrey[2], Philip Paty[1] and Francis Barany*

Department of Microbiology, Box 62, Hearst Microbiology Research Center, Strang Cancer Prevention Center, Joan and Sanford I. Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA, [1]Department of Colorectal Surgery, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA and [2]Department of Medicine, Joan and Sanford I. Weill Medical College of Cornell University, New York, NY 10021, USA

Received July 7, 1999; Revised October 8, 1999; Accepted October 18, 1999

ABSTRACT

Up to 15% of all colorectal cancers are considered to be replication error positive (RER$^+$) and contain mutations at hundreds of thousands of microsatellite repeat sequences. Recently, a number of intragenic mononucleotide repeat sequences have been demonstrated to be targets for inactivating genes in RER$^+$ colorectal tumors. In this study, thermostable DNA ligases were tested for the ability to detect alterations in microsatellite sequences in colon tumor samples. Ligation profiles on mononucleotide repeat sequences were determined for four related thermostable DNA ligases, *Thermus thermophilus* (*Tth*) ligase, *Thermus sp.* AK16D ligase, *Aquifex aeolicus* ligase and the K294R mutant of the *Tth* ligase. While the limit of detection for point mutations was one mutation in 1000 wild-type sequences, the ability to detect a single base deletion in a 10 base mononucleotide repeat was one mutation in 100 wild-type sequences. Furthermore, the misligation error increased exponentially as the length of the mononucleotide repeat increased, and was 10% of the correct signal for a 19 base mononucleotide repeat. A fluorescent ligase-based assay [polymerase chain reaction/ligase detection reaction (PCR/LDR)] correlated with results obtained using a radioactive assay to detect instability within the TGF-β Type II receptor gene. PCR/LDR was also used to detect the APCI1307K mononucleotide repeat allele which has a carrier frequency of 6.1% in Ashkenazi Jewish individuals. In a blind study, 30 samples that had been typed for the presence of the APCI1307K allele were tested. The PCR/LDR results correlated with those obtained using sequencing and allele-specific oligonucleotide hybridization for 16 samples carrying the mutation and 13 wild-type samples. Ligation assays that characterize mononucleotide repeats can be used to rapidly detect somatic mutations in tumors, and to screen for individuals who have a hereditary predisposition to develop colon cancer.

INTRODUCTION

Approximately 15% of the 150 000 new cases of colorectal cancer diagnosed each year demonstrate microsatellite instability (1,2). This instability is now understood to reflect the presence of the replication error positive (RER$^+$) phenotype in these tumors (3–8). The RER$^+$ phenotype results from an underlying molecular defect in the mismatch repair system which leads to a failure to correct frameshift errors in microsatellite repeats (9). Mutations in mismatch repair enzymes, hMSH2, hMLH1, hPMS1 and hPMS2, leading to RER$^+$ tumors, have been found in both hereditary cancers and spontaneous cancers (10–15). Over 95% of patients with hereditary non-polyposis colorectal cancer (HNPCC) syndrome have a predisposition for the development of RER$^+$ colorectal cancer due to inherited mutations in one of the mismatch repair enzymes (16,17).

Slippage errors during DNA replication can produce alterations in a microsatellite of defined length. In general, errors in multiple microsatellite repeats are diagnostic of the RER$^+$ phenotype (1–3). Furthermore, specific replication error mutations such as the deletion of an adenosine in an A10 repeat, have been reported in the TGF-β Type II receptor gene, providing a direct link between the RER$^+$ phenotype and a specific pathway of tumorigenesis (18). Analyses of RER$^+$ colorectal tumor samples indicate that >80% of such samples carry inactivating frameshift mutations in this poly(A) repeat, which represents an alternate pathway in the etiology of colorectal tumors (19,20).

The adenomatous polyposis coli (APC) gene represents another important cellular regulator in colonic epithelial cells. Disruption of APC function is an early event in colorectal tumorigenesis and is present in a majority of colon tumors (21–23). Germline mutations in the APC gene lead to an inherited colon cancer syndrome known as familial adenomatous polyposis syndrome (FAP), in which affected individuals present with thousands of polyps at a young age (24,25) Vogelstein and colleagues reported a polymorphism (T→A transversion at APC nucleotide 3920, codon 1307) found in 6.1% of Ashkenazi Jews and ~28% of Ashkenazim with a family

*To whom correspondence should be addressed. Tel: +1 212 746 6509; Fax: +1 212 746 8587; Email: barany@mail.med.cornell.edu

history of colon cancer (26). They have postulated that rather than altering the function of the encoded protein, this polymorphism creates a small hypermutable region of the gene, indirectly causing cancer predisposition through acquired somatic mutations. This transversion occurs within an interrupted stretch of seven poly(A) nucleosides. The polymorphism converts this region of the gene into an eight nucleotide long poly(A) microsatellite repeat. The somatic mutations observed in tumors from patients carrying this allele primarily include slippage within the microsatellite repeat, as well as some point mutations in the region (26).

The majority of mutation detection methods rely on the polymerase chain reaction (PCR) to initially amplify a target DNA region followed by subsequent analysis using either restriction endonuclease digestion, dideoxy sequencing, allele-specific hybridization, or polymerase-based detection (27). Microsatellites, in particular, are analyzed by PCR amplification followed by electrophoretic separation on gels (28,29). These methods are not well suited for the detection of variations occurring within mononucleotide repeats and closely clustered mutations, especially in tumor samples.

Our laboratory has developed a combined polymerase chain reaction/ligase detection reaction (PCR/LDR) method for discriminating single-base mutations (30–33). In LDR, thermostable DNA ligase specifically links two adjacent oligonucleotides when annealed at 65°C to a complementary target only when the nucleotides are perfectly base-paired at the junction. A single-base mismatch at the junction prevents ligation/amplification and is thus distinguished from a perfect match. While frameshift mutations occurring in mononucleotide repeats cannot be detected using allele-specific PCR, they may be amenable to detection using PCR/LDR. The feasibility of using PCR/LDR to detect the presence of a single nucleotide deletion in an A10 repeat sequence was tested using synthetic substrates. Subsequently, the assay was extended to detect alterations in mononucleotide repeat sequences in clinical samples. Finally, PCR/LDR was used to detect the APCI1307K polymorphism from DNA isolated from both blood and from archival paraffin-embedded colorectal tumor samples.

MATERIALS AND METHODS

Oligonucleotide synthesis and purification

Oligonucleotides were synthesized on an ABI 394 DNA Synthesizer (PE Biosystems Inc., Foster City, CA). Oligonucleotides used in LDR were purified by electrophoresis on 10% polyacrylamide/7 M urea gels. Bands were visualized by UV shadowing, excised from the gel, and eluted overnight at 64°C in TNE buffer (100 mM Tris–HCl pH 8.0, 500 mM NaCl, 5 mM EDTA). The oligonucleotides were recovered from the eluate using C18 Sep-Pak cartridges (Waters Corp., Milford, MA) following the manufacturer's instructions. Oligonucleotides were resuspended to ~1 mM in 100 µl TE (10 mM Tris pH 8.0, 1 mM EDTA). For LDR, gel purified stock solutions were diluted to 100 µM (100 pmol/µl).

The upstream, or discriminating oligonucleotides were labeled with fluorescent reporter groups at their 5′ termini during synthesis using Fam or Tet phosphoramidites (PE Biosystems Inc., Foster City, CA). The downstream, or common oligonucleotides, were phosphorylated at the 5′ end using a chemical phosphorylation reagent, and blocked at the 3′ end using a 3′-spacer C3 CPG (Glen Research, Sterling, VA).

DNA extraction from cell lines

Cell lines (SW620, HCT116 and DLD1) were grown in RPMI culture media with 10% fetal bovine serum. Harvested cells ($\sim 1 \times 10^7$) were resuspended in DNA extraction buffer (10 mM Tris–HCl pH 7.5, 150 mM NaCl, 2 mM EDTA) containing 0.5% SDS and 200 µg/ml proteinase K and incubated at 37°C for 4 h. Next, a 30% vol of 5 M NaCl was added and the mixture was centrifuged. DNA was precipitated from the supernatant with 3 vol of ethanol (EtOH), washed with 70% EtOH and resuspended in TE buffer (10 mM Tris–HCl pH 7.5, 2 mM EDTA pH 8.0).

DNA extraction from paraffin sections

DNA was prepared from 55 paraffin-embedded archival colon tumors and 36 paraffin-embedded pancreatic tumors. These samples represent a series of primary cancers removed by surgical resection at Memorial Sloan-Kettering Cancer Center. Tissue sections (10 sections, 10 µm thick) were cut from paraffin blocks and excess paraffin was removed. Samples were deparaffinized via sequential extraction with xylene, 100% EtOH and acetone, and dried under vacuum. The pellet was incubated overnight with proteinase K (200 µg/ml in 50 mM Tris–HCl pH 8.5, 1 mM EDTA and 0.5% Tween-20) at 55°C. After heating at 100°C for 10 min, debris was removed by centrifugation and the supernatant was stored at 4°C. Greater purity of DNA was achieved by phenol–chloroform extraction and ethanol precipitation, or by using the QIAamp Tissue Kit (Qiagen, Chatsworth, CA) when needed.

PCR conditions

PCR amplifications were carried out in a volume of 50 µl in 10 mM Tris–HCl buffer pH 8.3 containing 10 mM KCl, 4.0 mM $MgCl_2$, 250 µM dNTPs, 1 µM forward and reverse primers (50 pmol of each primer), and between 50 and 100 ng of genomic DNA extracted from paraffin blocks or from cell lines. After a 10 min denaturation step, 1.5 U of *Amplitaq Gold* DNA polymerase (Perkin Elmer, Norwalk, CT) was added under hot start conditions, and amplification achieved by thermally cycling for 35 or 40 cycles of 94°C for 30 s, 60°C for 1 min, 72°C for 1 min and 72°C for 3 min for a final extension. Four microliters of the PCR product were analyzed on a 2% agarose gel to verify presence of amplification product of the expected size. For DNA isolated from paraffin sections, *Pfu* polymerase (Stratagene, La Jolla, CA) was used for PCR amplification following the manufacturer's instructions. The sequences of PCR primers used are listed in Table 1.

LDR conditions

LDR reactions were carried out in a 20 µl mixture containing 20 mM Tris–HCl pH 7.6, 10 mM $MgCl_2$, 100 mM KCl, 10 mM DTT, 1 mM NAD^+, 25 nM (500 fmol) of the detecting primers and mixtures of PCR products from cell lines or tumor samples. The reaction mixture was heated for 1.5 min at 94°C prior to adding 25 fmol of the wild-type or mutant *Thermus thermophilus* (*Tth*) DNA ligase. Ligases were overproduced and purified as described previously (34). *Tth* DNA ligase is available from New England Biolabs. LDR reactions were

Table 1. PCR and LDR primers used to analyze mutations in intragenic mononucleotide repeats in the TGF-β Type II receptor and APC genes

Primer type	Primer sequence
TGF forward	TGF Top1 T_m: 70.1°C
PCR primer	5′-CTT TAT TCT GGA AGA TGC TGC TTC TC-3′
TGF reverse	TGF Bot2R T_m: 69.6°C
PCR primer	5′-GTC AGG ATT GCT GGT GTT ATA TTC TTC-3′
APC forward	APC1307 Top1 T_m = 70.6°C
PCR primer	5′-GCT GCC ACT TGC AAA GTT TCT TC-3′
APC reverse	APC1307 Bot2R T_m = 71.5°C
PCR primer	5′-AGT GGG GTC TCC TGA ACA TAG TGT TC-3′
TGF common	TGF common T_m = 66.3°C
LDR primer	5′-p AAA AGC CTG GTG AGA CTT TCT TC-3′
TGF mutant	Fam-TGF mt5A T_m = 68.5°C
LDR primer	5′-Fl- AC CAA AGT GCA TTA TGA AGG AAA AA-3′
TGF wildtype	Fam-TGF wt6A T_m = 66.5°C
LDR primer	5′-Fl- AA ACA AAG TGC ATT ATG AAG GAA AAA A-3′
APC common	APC1307 common T_m = 67.0°C
LDR primer	5′-p AAA AGA AAA GAT TGG AAC TAG GTC AG-3′
APC wild-type	APC1307 wt3AT T_m = 66.6°C
LDR primer	5′-Fl- CTA ATA CCC TGC AAA TAG CAG AAA T-3′
APC mutant	APC1307 mt4A T_m = 66.6°C
LDR primer	5′-Fl- A AAA CTA ATA CCC TGC AAA TAG CAG AAA A-3′

thermally cycled for 20 cycles of 15 s at 94°C and 2 min at 65°C. Reactions were stopped by adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% formamide, 8.3 mM EDTA and 0.17% Blue Dextran). The mixture was supplemented with 0.5 µl Rox-1000, or TAMRA 350 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea–10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 V for 2 h. Fluorescent ligation products were analyzed and quantified using the ABI Gene Scan 672 software. The amount of product is calculated from a calibration curve (one fmol = 600 peak area units). The sequences of LDR primers used are listed in Table 1.

Sequencing

PCR products were sequenced using fluorescent dideoxy nucleotides and *Taq* FS polymerase (Perkin Elmer Cetus, Norwalk, CT). Following an initial denaturation at 94°C for 2 min, the sequencing reactions were carried out using 25 cycles of 94°C for 10 s, 50°C for 5 s, 60°C for 3 min. The products were electrophoresed as described above and analyzed using ABI Genescan sequence analysis software.

Restriction analysis

PCR primers used to convert the APC1307 region to an *Mbo*II site or an *Ase*I site were 5′-CTG CTA ATA CCC TGC AAA TAG CAG AA G-3′ and 5′-AGA TTC TGC TAA TAC CCT GCA AAT AGC ATT AA-3′, respectively. The reverse primer used for PCR in both cases was 5′-TGC TGT GAC ACT GCT GGA ACT TC-3′. Aliquots of PCR reactions were digested using either *Mbo*II or *Ase*I (New England Biolabs, Beverly, MA) in the appropriate buffers. The digested products were electrophoresed on a 3% agarose gel, and stained with ethidium bromide.

RESULTS AND DISCUSSION

Detection of mononucleotide repeat polymorphisms using thermostable ligase

A preliminary assay was devised to determine whether a ligase-based assay could detect slippage mutations in a mononucleotide microsatellite repeat sequence. The region of the TGF-β Type II receptor region containing a mononucleotide repeat sequence was synthesized and used to test the specificity of *Tth* ligase in discriminating a mononucleotide repeat of A10 from A9 and A8 (Fig. 1, top). NAD^+-dependent thermostable ligases cloned from *Aquifex aeolicus*, *T.thermophilus* and *Thermus AK16D*, were compared using this ligation assay. All of these thermostable DNA ligases demonstrated high fidelity in an LDR assay for frameshift mutations (results for *Tth* ligase are shown in Fig. 1, bottom).

In order to test whether the position of the ligation junction along a repetitive sequence would affect ligation fidelity, ligation primers which positioned the ligation junction at different points along the mononucleotide repeat sequence were synthesized. Templates containing either nine or 10 mononucleotide repeats

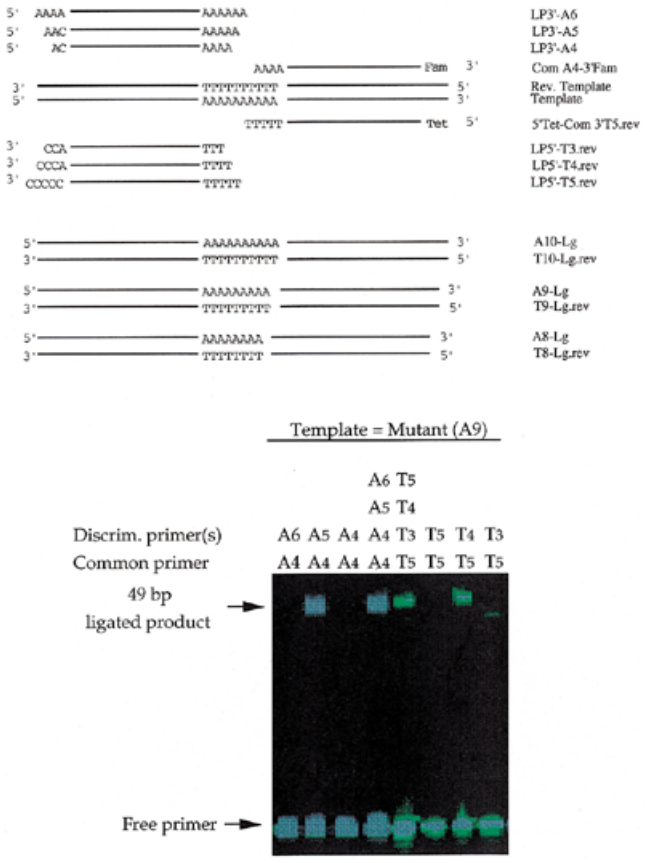

Figure 1. (Top) Synthetic substrate assay designed to test LDR on mononucleotide repeat sequences. Primers designed to detect A10, A9 and A8 templates. Primer Com A4 is phosphorylated on its 5′ terminus and has a Fam-label on its 3′ terminus. The primers LP3′-A6, LP3′-A5 and LP3′-A4, have repeat sequences at their 3′ termini and varying length tails at their 5′ termini. Primer 5′Tet-Com-3T5.rev has a Tet-label on its 5′ terminus and five thymidines on its 3′ terminus. The primers LP5′-T3.rev, LP5′-T4.rev and LP5′-T5.rev are phosphorylated on their 5′ termini and have varying length tails on their 3′ termini. The ligase reaction was carried out using 500 fmol of each primer and 50 fmol of template DNA in a 20 µl reaction with 25 fmol of *Tth* ligase per reaction. After ligation, the primers were separated and analyzed on an ABI 373 automated DNA sequencer. (Bottom) The primers designed for the poly(A) target were tested using the LDR on a template containing an A9 repeat. The ligation reaction could determine the length of the template in either a uniplex (side lanes) or multiplex format (center two lanes). There was only minimal misligation in the last lane.

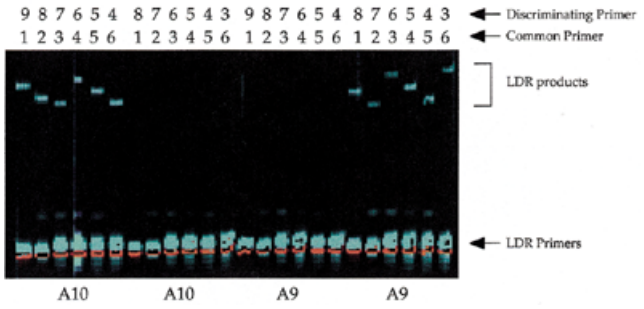

Figure 2. LDR results for varying the position of the ligation junction through a repeat region. The ligation reaction was tested using common primers that contained from one to six adenosines at their 3′ termini (lower set of numbers above gel) which were paired with discriminating primers that contained the remaining adenosines required to form an A10 or A9 repeat (upper set of numbers above gel). The results of ligation are shown for both A10 and A9 templates.

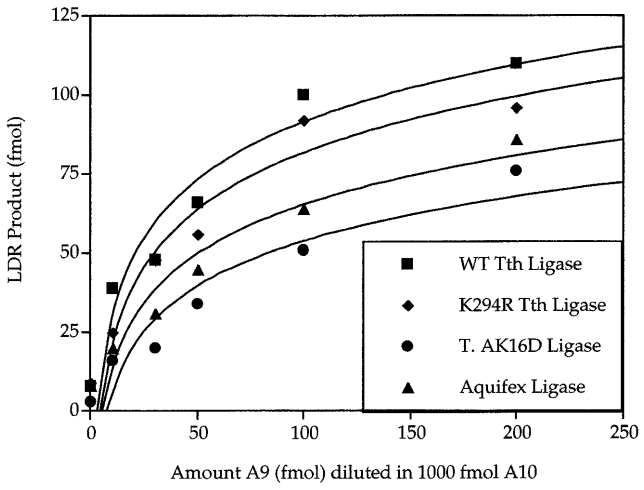

Figure 3. Ligation profiles of the wild-type *Tth*, K294R *Tth*, *Thermus AK16D* and *Aquifex* DNA ligases for different concentrations of target DNA template. Various concentrations of A9 template (from 0 to 200 fmol) were diluted into 1000 fmol of A10 template. Ligase detection reactions were carried out as described in Materials and Methods. The graph depicts the amount of LDR product formed for each enzyme.

generated ligation products only in the presence of oligonucleotides that perfectly matched the number of repeated nucleotides. The site of the ligation junction along the nucleotide repeat did not influence the ability of the oligonucleotides to ligate. Thus, for short mononucleotide repeats, the ligation fidelity was determined to be independent of position of the ligation junction along the mononucleotide repeat sequence (Fig. 2). A study of steady-state kinetics for discrimination of a G:T mismatch demonstrated that a K294R mutant *Tth* ligase, *Tsp AK16D* ligase and *Aquifex* ligase demonstrated greater ligation fidelity compared to the wild-type *Tth* DNA ligase (35,36). A comparison of these four ligases on a mononucleotide target sequence yielded similar ligation profiles for each enzyme. These enzymes could detect the presence of a single-base frameshift mutation in up to a 100-fold excess of wild-type DNA sequence (Fig. 3). Calibration curves were constructed for the *Tth* DNA ligase enzyme to compare fidelity for point mutations versus frameshift mutations in a repetitive sequence (Fig. 4). LDR could detect one point mutation in 1000 copies of wild-type sequence and was quantitative over a three-log range (33). However, for frameshift mutations in a repetitive region, LDR was quantitative for less than a two-log range and could detect one mutation in 100 copies of wild-type sequence. On the other hand, the *Pyrococcus* and T4 ligases could not detect single-base frameshift mutations within mononucleotide repeats due to the high level of background misligation on wild-type sequences (data not shown). This difference was primarily a result of the higher rate of misligation for the frameshift mutations, as compared to the misligation rate for point mutations.

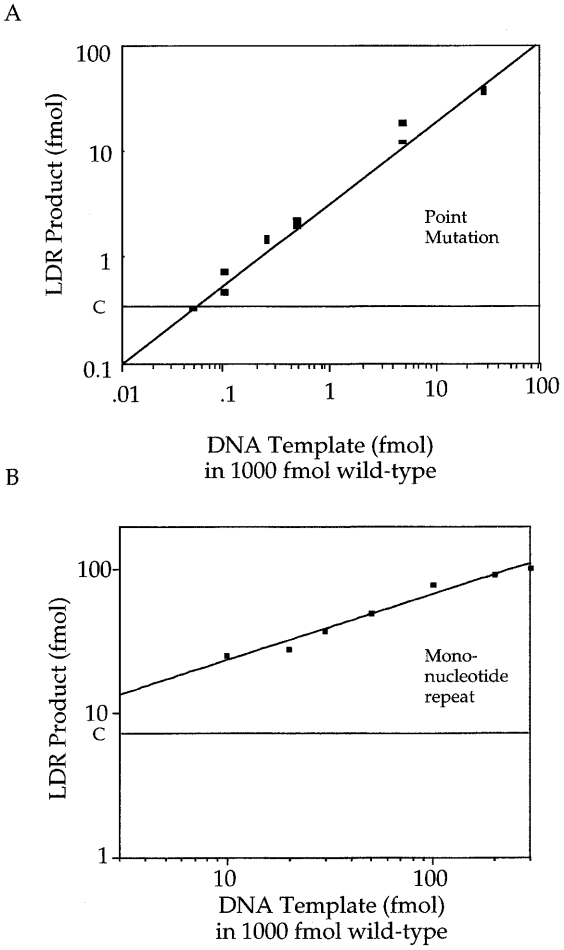

Figure 4. Calibration curves for LDR for point mutations and frameshift mutations using *Tth* DNA ligase. The 'C' point represents the amount of noise generated from misligation of the LDR primers on the excess wild-type template present in the reaction. The ligase detection reaction is quantitative over three orders of magnitude for point mutations (**A**), but less than two orders of magnitude for frameshift mutations in repetitive regions (**B**). In both cases, the ligase detection reactions were carried out using a dilution of mutant template in 1000 fmol of wild-type template. The detection limit is one mutant in 1000 wild-type sequences for point mutations and one mutant in 100 wild-type sequences for slippage mutations in repetitive sequences.

As the length of mononucleotide repeats is increased, a sequence will be reached for which the *Tth* DNA ligase would tolerate single-base insertions or deletions. To determine this length, templates containing longer poly(A) stretches were synthesized, gel-purified and tested. In an LDR assay, the *Tth* DNA ligase was able to distinguish an A16 template from an A15 template. There was only a slight increase in the amount of product generated from misligations for the incorrect template, ranging up to 4% of the correct product signal. For LDR on templates up to 19 mononucleotide repeats, the amount of signal generated from incorrect hybridization of the LDR primers to the synthetic templates increased to ~10% of the correct product signal.

PCR amplification of mononucleotide repeat sequences

The greatest source of error for analysis of mononucleotide repeat sequences, however, is the error generated during PCR amplification of microsatellite repeats. Polymerases that lack exonucleolytic activity typically generate non-templated 3′-terminal nucleotides that interfere with proper sizing of PCR amplicons. In addition, polymerases that possess exonuclease activity remove 3′-terminal bases from PCR amplicons and therefore produced a mixed population of PCR products (37). In either case, determination of mononucleotide repeat length based solely on the size of PCR products is complicated by these factors. In contrast, neither of these modifications of PCR products affect the ability to use LDR to determine the size of a mononucleotide repeat in a PCR product. Accumulation of errors during PCR amplification usually compounds the error rate of the DNA polymerase utilized (38). Therefore, even a moderate increase in replication fidelity of a polymerase can produce significant improvements in the fidelity of PCR amplification

A variety of PCR conditions were tested for genomic DNA samples in order to minimize the amount of slippage in the repeat regions during PCR amplification. The enzymes tested included *Taq* and *AmpliTaq Gold* polymerase, Vent (exo^+) polymerase, Vent (exo^-) polymerase, UlTma polymerase and *Pfu* polymerase. Each of the polymerases was tested in its native buffers with 2 mM Mg^{2+} and a dNTP concentration of 250 μM. Of the polymerases tested, *Taq* demonstrated the least fidelity, generating 50% slippage error when amplifying a control plasmid containing the TGF-β Type II receptor A10 repeat sequence. The polymerase with the greatest fidelity in amplification of this A10 region was *Pfu* polymerase, with an error band that was <20% of the control band after 30 rounds of PCR amplification. However, for amplification of poly-guanosine or polycytosine mononucleotide repeats (i.e., the C8 region of the BAX gene), *AmpliTaq* polymerase did not generate slippage artifacts (data not shown).

Detection of mononucleotide repeat mutations in TGF-β Type II receptor using PCR/LDR

The HCT116 cell line has been reported in the literature to be RER^+ and demonstrates slippage in the TGF-β Type II receptor (18,39,40). Using optimized amplification conditions, genomic DNA extracted from the HCT116 cell line and human umbilical cord blood were typed to verify the presence and absence of mononucleotide slippage mutations. Subsequently, 12 carotid endarterectomy specimens (40) were typed using PCR/LDR and the results were compared to those obtained by PCR amplification followed by restriction digestion of the PCR amplicon to generate DNA fragments that could be sized precisely. The results were concordant using both methods for all of the samples tested and there was quantitative correlation between the results in each case (Fig. 5).

Detection of the APCI1307K mononucleotide repeat polymorphism in clinical samples using PCR/LDR

The APCI1307K allele creates an A8 stretch within the coding region of the APC gene. LDR primers were designed for the detection of the APCI1307K allele and were initially tested on synthetic substrates. Blinded genomic DNA samples, initially typed using allele specific oligonucleotides (ASOs) were obtained from Dr Ostrer's laboratory at the Human Genetics Program at New York University Medical College. All positive samples and some negative samples had been sequenced to confirm the initial result in these samples (26). Thirty samples

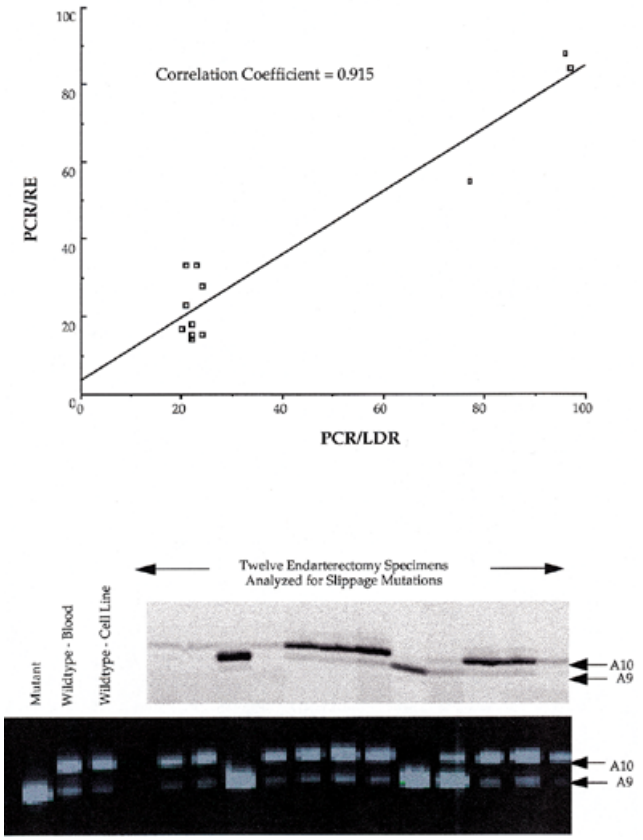

Figure 5. Comparison of PCR/LDR with PCR/RE for analysis of TGF-β Type II receptor slippage mutations. Genomic DNA from carotid endarterectomy specimens were analyzed for mutations in the TGF-β Type II receptor by primary PCR amplification followed by either restriction digestion to remove PCR sizing artifacts or LDR. The relative percentage of mutant compared to wild-type signals were quantitated using a phosphoimager for PCR/RE and using an ABI 373 for PCR/LDR. (Top) Graph depicting the correlation between the two methods. (Bottom) Gel results from PCR/RE and PCR/LDR analysis of 12 carotid endarterectomy specimens. The samples included two mutant samples and one sample which was a primary culture of a mutant sample that was partially transfected with a plasmid containing wild-type TGF-β Type II receptor.

were tested using PCR/LDR in an initial study and the same 30 samples were resorted and retested in a follow-up study. Representative PCR/LDR results for detection of the APCI1307K allele are shown (Fig. 6, top). In both studies, the results obtained using the PCR/LDR method matched those obtained by the hybridization and sequencing except for one sample. This sample was repeatedly typed as positive by PCR/LDR and had been typed as negative by ASO hybridization. The sample was sequenced using fluorescent dideoxy sequencing and the presence of the APCI1307K allele was confirmed. This sample therefore represents a false negative missed by the hybridization method and correctly typed using the PCR/LDR primers.

In a subsequent study to determine whether PCR/LDR could be used to screen for the APCI1307K allele in archival specimens, genomic DNA obtained from paraffin-embedded tissue was analyzed (33). In this study, archival tumor specimens from 55 Ashkenazi Jewish patients treated at Memorial Sloan-Kettering Cancer Center were analyzed (41). Genomic DNA extracted from these samples was analyzed using PCR/LDR, and seven of the 55 samples were found to carry the APCI1307K allele. Four of the positive samples found by PCR/LDR analysis are shown (Fig. 6, bottom). In contrast, only one of the 36 pancreatic cancer samples tested positive for the APCI1307K allele. The prevalence rate of the allele among the colorectal cancer samples (12.7%) was similar to the rate published in the original Vogelstein study (10.4%) (26). This rate is significantly higher than the 6.1% carrier frequency of the polymorphism in the Ashkenazi Jewish population ($\chi^2 = 4.4$, $P < 0.036$). This result is in agreement with the statistically significant association obtained between the development of colorectal cancer and the APCI1307K allele in the study by Laken *et al.* (26).

Comparison of amplification-created restriction site (ACRS) analysis and PCR/LDR for detecting the APCI1307K mononucleotide repeat polymorphism

The seven APCI1307K$^+$ tumor samples were retested using PCR/LDR to confirm the presence of the APCI1307K allele. The samples were also sequenced and the presence of the allele was confirmed in the seven samples. ACRS analysis was also used in order to confirm the presence of the APCI1307K allele. The restriction enzyme used was *Ase*I, which requires two base substitutions to convert the wild-type APC1307 sequence (AGAAAT) to its ATTAAT recognition sequence. The *Ase*I conversion leaves the APCI1307K allele undigested, but digests sequences that were originally wild-type. Using the *Ase*I conversion method, the presence of the APCI1307K allele was confirmed for the seven APCI1307K+ tumor samples, although PCR amplified fragments resistant to *Ase*I digestion were also observed in wild-type samples (data not shown). Attempts to create an *Mse*I restriction site (AAAT to AATT) gave false positive results, the consequence of polymerase misextension from a 3′-terminal mismatched base (42,43).

CONCLUDING REMARKS

PCR/LDR has been shown to be a powerful tool in cancer and disease gene mutation detection (30–33,44). In this study, the specificity and fidelity limits of different thermostable ligase enzymes were compared for the analysis of polymorphisms in mononucleotide repeat sequences of different lengths. PCR conditions with various thermostable polymerases were compared in order to create a PCR/LDR assay for characterizing microsatellite instability in clinical samples. The optimal PCR conditions for amplification of polyadenosine mononucleotide repeat DNA necessitated the use of thermostable polymerases with exonucleolytic proofreading activity.

The genetic alterations found in RER$^+$ colorectal tumors occur primarily in repetitive DNA sequences such as mononucleotide repeats. RER$^+$ tumors are less likely to have mutations in p53, are poorly differentiated, and are more prone to metastasize to other organs (45). Approximately 89% of patients with multiple primary tumors have RER$^+$ tumors (46).

Genes containing intragenic mononucleotide repeats which have been shown to be altered in RER$^+$ colorectal tumors include the TGF-β Type II receptor, Bax gene, hMSH3 and hMSH6 genes (1,2,18,47–53). These intragenic mononucleotide repeats are all amenable to analysis using PCR/LDR (unpublished results). Furthermore, screening for the APCI1307K polymorphism

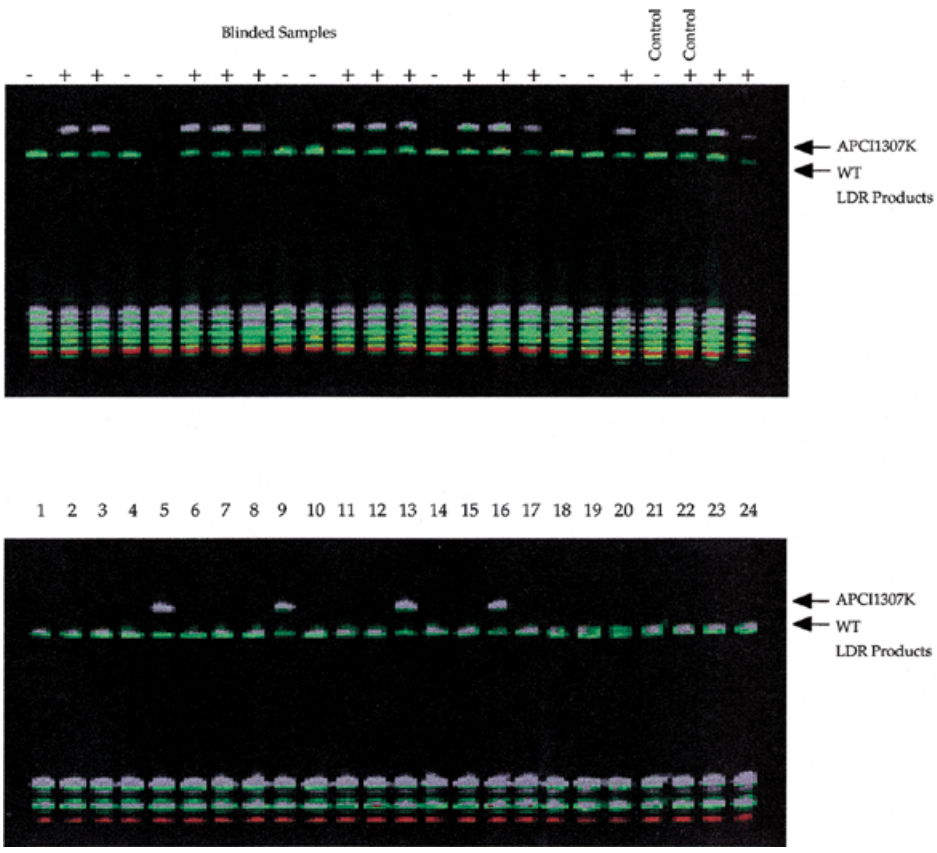

Figure 6. (Top) Detection of the APCI1307K allele in genomic DNA isolated from blood samples of Ashkenazi Jewish individuals using PCR/LDR. The upper band in each lane represents the mutant APCI1307K allele and the lower band represents the wild-type allele. In each case, the results were concordant with results obtained using hybridization with ASOs. +, mutation present; –, mutation absent. (Bottom) Detection of the APCI1307K allele in paraffin-embedded archival colorectal tumors. Genomic DNA was extracted from paraffin-embedded colorectal tumors from Ashkenazi Jewish individuals. The DNA was analyzed for the APCI1307K allele using PCR/LDR. The gel shows the results for 24 samples in which samples 5, 9, 13 and 16 were found to carry the APCI1307K allele.

was demonstrated on both genomic DNA from whole blood and archival tumor specimens from Ashkenazi Jewish individuals.

Multiplexed ligation reactions have now been validated to distinguish 19 possible mutations in K-*ras* (33), over 30 mutations in the cystic fibrosis transmembrane conductance regulator gene (44), and over 55 mutations in the *p53* gene (unpublished results). PCR/LDR reactions are compatible with a universal DNA array detection scheme (54), and with this work, we extend this versatility to include mononucleotide repeat sequences as well.

This level of multiplexing is not possible with other mutation detection methods such as allele-specific PCR (AS-PCR), ASO hybridization, dideoxy sequencing, or restriction endonuclease digestion. For AS-PCR, attempts to analyze more than four mutations have required splitting different sets of primers into different tubes either because of closely clustered mutations or failure to PCR amplify all allele-specific primers simultaneously (55–61). Likewise, restriction digestion methods are not amenable to multiplexing. Further, for both allele-specific and ACRS techniques, false positive results may be obtained due to polymerase misextension from a 3′-terminal mismatched base (42,43).

Molecular characterization of colorectal tumors may eventually play a role in the clinical management of this disease. There is already some indication that adjuvant chemotherapy based on the status of molecular markers may be beneficial for some patients (62). Microsatellite instability in colorectal tumors is associated with diploid status, tumors in the proximal colon, and poor differentiation (63). Preliminary studies have also shown that patients with tumors that demonstrated microsatellite instability at two or more loci had an increased overall survival rate when compared to patients with tumors which lacked genomic instability (64). The ability to rapidly characterize alterations in microsatellite repeats using ligase-based assays may eventually aid in the clinical management of colorectal cancer.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Harry Ostrer and Carol Oddoux from the Human Genetic Program at New York University Medical College for providing blind genomic DNA samples previously typed for the APCI1307K allele from Ashkenazi Jewish individuals for the APCI1307K experiments. The authors would also like to acknowledge Norman Gerry, Reyna Favis, Weiguo Cao, Jie Tong, Joseph Day, Andrew Grace and Matt D'Alessio for their helpful suggestions and technical assistance. M.Z. was supported by the William M. Keck Fellowship and the Tri-Institutional Medical Scientist Training

Program (MSTP) Fellowship during the course of this study. Support for this work was provided by the National Cancer Institute (P01-CA65930 and RO1-CA81467). T.M. and G.N. were supported by NHLBI-SCOR HL56987 and NIA AG12712.

REFERENCES

1. Ionov,Y., Peinado,M.A., Malkhosyan,S., Shibata,D. and Perucho,M. (1993) *Nature*, **363**, 558–561.
2. Shibata,D., Peinado,M.A., Ionov,Y., Malkhosyan,S. and Perucho,M. (1994) *Nat. Genet.*, **6**, 273–281.
3. Chen,J., Heerdt,B.G. and Augenlicht,L.H. (1995) *Cancer Res.*, **55**, 174–180.
4. Cawkwell,L., Lewis,F.A. and Quirke,P. (1994) *Br. J. Cancer*, **70**, 813–818.
5. Liu,B., Nicolaides,N.C., Markowitz,S., Willson,J.K., Parsons,R.E., Jen,J., Papadopolous,N., Peltomaki,P., de la Chapelle,A., Hamilton,S.R. *et al.* (1995) *Nat. Genet.*, **9**, 48–55.
6. Orth,K., Hung,J., Gazdar,A., Bowcock,A., Mathis,J.M. and Sambrook,J. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 9495–9499.
7. Aaltonen,L.A., Peltomaki,P., Leach,F.S., Sistonen,P., Pylkkanen,L., Mecklin,J.P., Jarvinen,H., Powell,S.M., Jen,J., Hamilton,S.R. *et al.* (1993) *Science*, **260**, 812–816.
8. Risinger,J.I., Berchuck,A., Kohler,M.F., Watson,P., Lynch,H.T. and Boyd,J. (1993) *Cancer Res.*, **53**, 5100–5103.
9. Perucho,M. (1996) *Nature Med.*, **2**, 630–631.
10. Fishel,R., Lescoe,M.K., Rao,M.R.S., Copeland,N.G., Jenkins,N.A., Garber,J., Kane,M. and Kolodner,R. (1993) *Cell*, **75**, 1027–1038.
11. Parsons,R., Li,G.M., Longley,M.J., Fang,W.H., Papadopoulos,N., Jen,J., de la Chapelle,A., Kinzler,K.W., Vogelstein,B. and Modrich,P. (1993) *Cell*, **75**, 1227–1236.
12. Papadopoulos,N., Nicolaides,N.C., Wei,Y.F., Ruben,S.M., Carter,K.C., Rosen,C.A., Haseltine,W.A., Fleischmann,R.D., Fraser,C.M., Adams,M.D. *et al.* (1994) *Science*, **263**, 1625–1629.
13. Leach,F.S., Nicolaides,N.C., Papadopoulos,N., Liu,B., Jen,J., Parsons,R., Peltomaki,P., Sistonen,P., Aaltonen,L.A., Nystrom-Lahti,M. *et al.* (1993) *Cell*, **75**, 1215–1225.
14. Nicolaides,N.C., Papadopoulos,N., Liu,B., Wei,Y.F., Carter,K.C., Ruben,S.M., Rosen,C.A., Haseltine,W.A., Fleischmann,R.D., Fraser,C.M. *et al.* (1994) *Nature*, **371**, 75–80.
15. Bronner,C.E., Baker,S.M., Morrison,P.T., Warren,G., Smith,L.G., Lescoe,M.K., Kane,M., Earabino,C., Lipford,J., Lindblom,A. *et al.* (1994) *Nature*, **368**, 258–261.
16. Lynch,H.T. and Smyrk,T. (1996) *Cancer*, **78**, 1149–1167.
17. Brassett,C., Joyce,J.A., Froggatt,N.J., Williams,G., Furniss,D., Walsh,S., Miller,R., Evans,D.G. and Maher,E.R. (1996) *J. Med. Genet.*, **33**, 981–985.
18. Markowitz,S., Wang,J., Myeroff,L., Parsons,R., Sun,L., Lutterbaugh,J., Fan,R.S., Zborowska,E., Kinzler,K.W., Vogelstein,B. *et al.* (1995) *Science*, **268**, 1336–1338.
19. Parsons,R., Myeroff,L.L., Liu,B., Willson,J.K., Markowitz,S.D., Kinzler,K.W. and Vogelstein,B. (1995) *Cancer Res.*, **55**, 5548–5550.
20. Brattain,M.G., Markowitz,S.D. and Willson,J.K. (1996) *Curr. Opin. Oncology*, **8**, 49–53.
21. Ilyas,M. and Tomlinson,I.P. (1997) *J. Pathol.*, **182**, 128–137.
22. Powell,S.M., Zilz,N., Beazer-Barclay,Y., Bryan,T.M., Hamilton,S.R., Thibodeau,S.N., Vogelstein,B. and Kinzler,K.W. (1992) *Nature*, **359**, 235–237.
23. Kinzler,K.W. and Vogelstein,B. (1996) *Cell*, **87**, 159–170.
24. Fearon,E.R. and Vogelstein,B. (1990) *Cell*, **61**, 759–767.
25. Eckert,W.A., Jung,C. and Wolff,G. (1994) *J. Med. Genet.*, **31**, 442–447.
26. Laken,S.J., Petersen,G.M., Gruber,S.B., Oddoux,C., Ostrer,H., Giardiello,F.M., Hamilton,S.R., Hampel,H., Markowitz,A., Klimstra,D., Jhanwar,S., Winawer,S., Offit,K., Luce,M.C., Kinzler,K.W. and Vogelstein,B. (1997) *Nat. Genet.*, **17**, 79–83.
27. Weber,J.L. (1990) *Curr. Opin. Biotechnol.*, **1**, 166–171.
28. Mansfield,D.C., Brown,A.F., Green,D.K., Carothers,A.D., Morris,S.W., Evans,H.J. and Wright,A.F. (1994) *Genomics*, **24**, 225–233.
29. Ziegle,J.S., Su,Y., Corcoran,K.P., Nie,L., Mayrand,P.E., Hoff,L.B., McBride,L.J., Kronick,M.N. and Diehl,S.R. (1992) *Genomics*, **14**, 1026–1031.
30. Barany,F. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 189–193.
31. Barany,F. (1991) *PCR Methods Appl.*, **1**, 5–16.
32. Day,D., Speiser,P.W., White,P.C. and Barany,F. (1995) *Genomics*, **29**, 152–162.
33. Khanna,M., Park,P., Zirvi,M., Paty,P. and Barany,F. (1999) *Oncogene*, **18**, 27–38.
34. Barany,F. and Gelfand,D. (1991) *Gene*, **109**, 1–11.
35. Luo,J., Bergstrom,D.E. and Barany,F. (1996) *Nucleic Acids Res.*, **24**, 3071–3078.
36. Tong,J., Cao,W. and Barany,F. (1999) *Nucleic Acids Res.*, **27**, 788–794.
37. Hite,J.M., Eckert,K.A. and Cheng,K.C. (1996) *Nucleic Acids Res.*, **24**, 2429–2434.
38. Eckert,K.A. and Kunkel,T.A. (1991) *PCR Methods Appl.*, **1**, 17–24.
39. Casares,S., Ionov,Y., Ge,H.Y., Stanbridge,E. and Perucho,M. (1995) *Oncogene*, **11**, 2303–2310.
40. McCaffrey,T.A., Du,B., Consigli,S., Szabo,P., Bray,P.J., Hartner,L., Weksler,B.B., Sanborn,T.A., Bergman,G. and Bush,H.L. (1997) *J. Clin. Invest.*, **100**, 2182–2188.
41. Guillem,J.G., Paty,P.B. and Cohen,A.M. (1997) *Cancer J. Clin.*, **47**, 113–128.
42. Day,J., Bergstrom,D., Hammer,R. and Barany,F. (1999) *Nucleic Acids Res.*, **27**, 1810–1818.
43. Day,J., Hammer,R., Bergstrom,D. and Barany,F. (1999) *Nucleic Acids Res.*, **27**, 1819–1827.
44. Eggerding,F.A., Iovannisci,D.M., Brinson,E., Grossman,P. and Winn-Deen,E.S. (1995) *Hum. Mutat.*, **5**, 153–165.
45. Muta,H., Noguchi,M., Perucho,M., Ushio,K., Sugihara,K., Ochiai,A., Nawata,H. and Hirohashi,S. (1996) *Cancer*, **77**, 265–270.
46. Horii,A., Han,H.J., Shimada,M., Yanagisawa,A., Kato,Y., Ohta,H., Yasui,W., Tahara,E. and Nakamura,Y. (1994) *Cancer Res.*, **54**, 3373–3375.
47. Simms,L.A., Zou,T.T., Young,J., Shi,Y.Q., Lei,J., Appel,R., Rhyu,M.G., Sugimura,H., Chenevix-Trench,G., Souza,R.F., Meltzer,S.J. and Leggett,B.A. (1997) *Oncogene*, **14**, 2613–2618.
48. Ouyang,H., Shiwaku,H.O., Hagiwara,H., Miura,K., Abe,T., Kato,Y., Ohtani,H., Shiiba,K., Souza,R.F., Meltzer,S.J. and Horii,A. (1997) *Cancer Res.*, **57**, 1851–1854.
49. Lengauer,C., Kinzler,K.W. and Vogelstein,B. (1997) *Nature*, **386**, 623–627.
50. Rampino,N., Yamamoto,H., Ionov,Y., Li,Y., Sawai,H., Reed,J.C. and Perucho,M. (1997) *Science*, **275**, 967–969.
51. Souza,R.F., Appel,R., Yin,J., Wang,S., Smolinski,K.N., Abraham,J.M., Zou,T.T., Shi,Y.Q., Lei,J., Cottrell,J., Cymes,K., Biden,K., Simms,L., Leggett,B., Lynch,P.M., Frazier,M., Powell,S.M., Harpaz,N., Sugimura,H., Young,J. and Meltzer,S.J. (1996) *Nat. Genet.*, **14**, 255–257.
52. Yamamoto,H., Sawai,H. and Perucho,M. (1997) *Cancer Res.*, **57**, 4420–4426.
53. Yamamoto,H., Sawai,H., Weber,T.K., Rodriguez-Bigas,M.A. and Perucho,M. (1998) *Cancer Res.*, **58**, 997–1003.
54. Gerry,N., Witowski,N., Barany,G. and Barany,F. (1999) *J. Mol. Biol.*, **292**, 251–262.
55. Wenham,P.R., Newton,C.R. and Price,W.H. (1991) *Clin. Chem.*, **37**, 241–244.
56. Zschocke,J. and Graham,C.A. (1995) *Mol. Cell. Probes*, **9**, 447–451.
57. Fortina,P., Dotti,G., Conant,R., Monokian,G., Parrella,T., Hitchcock,W., Rappaport,E., Schwartz,E. and Surrey,S. (1992) *PCR Methods Appl.*, **2**, 163–166.
58. Baty,D., Terron Kwiatkowski,A., Mechan,D., Harris,A., Pippard,M.J. and Goudie,D. (1998) *J. Clin. Pathol.*, **51**, 73–74.
59. Scobie,G., Woodroffe,B., Fishel,S. and Kalsheker,N. (1996) *Mol. Hum. Reprod.*, **2**, 203–207.
60. Patel,P., Lo,Y.M., Bell,J.I. and Wainscoat,J.S. (1993) *J. Clin. Pathol.*, **46**, 1105–1108.
61. Ferrie,R.M., Schwarz,M.J., Robertson,N.H., Vaudin,S., Super,M., Malone,G. and Little,S. (1992) *Am. J. Hum. Genet.*, **51**, 251–262.
62. O'Connell,M.J., Schaid,D.J., Ganju,V., Cunningham,J., Kovach,J.S. and Thibodeau,S.N. (1992) *Cancer*, **70** (Suppl), 1732–1739.
63. Lothe,R.A., Peltomaki,P., Meling,G.I., Aaltonen,L.A., Nystrom-Lahti,M., Pylkkanen,L., Heimdal,K., Andersen,T.I., Moller,P., Rognum,T.O. *et al.* (1993) *Cancer Res.*, **53**, 5849–5852.
64. Sidransky,D. (1997) *Science*, **278**, 1054–1059.