Oncogene (2002) 21, 1909–1921

www.nature.com/onc

TECHNICAL REPORT

An endonuclease/ligase based mutation scanning method especially suited for analysis of neoplastic tissue

Jianmin Huang[1], Brian Kirk[1], Reyna Favis[1], Thierry Soussi[2], Philip Paty[3], Weiguo Cao[1,4] and Francis Barany*[,1]

[1]*Department of Microbiology, Box 62, Hearst Microbiology Research Center, Strang Cancer Prevention Center, Joan and Sanford I Weill Medical College of Cornell University, Room B-406, 1300 York Avenue, New York, NY 10021, USA;* [2]*Laboratoire de Génotoxicologie des tumeurs, Institut Curie, 26 rue d'Ulm, 75248 Paris, France;* [3]*Colorectal Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, Room C-983, 1275 York Avenue, New York, NY 10021, USA*

Knowledge of inherited and sporadic mutations in known and candidate cancer genes may influence clinical decisions. We have developed a mutation scanning method that combines thermostable EndonucleaseV (Endo V) and DNA ligase. Variant and wild-type PCR amplicons are generated using fluorescently labeled primers, and heteroduplexed. *Thermotoga maritima (Tma)* EndoV recognizes and primarily cleaves heteroduplex DNA one base 3′ to the mismatch, as well as nicking matched DNA at low levels. *Thermus* species (*Tsp.*) AK16D DNA ligase reseals the background nicks to create a highly sensitive and specific assay. The fragment mobility on a DNA sequencing gel reveals the approximate position of the mutation. This method identified 31/35 and 8/8 unique point mutations and insertions/deletions, respectively, in the *p53*, *VHL*, K-*ras*, *APC*, *BRCA1*, and *BRCA2* genes. The method has the sensitivity to detect K-*ras* mutations diluted 1 : 20 with wild-type DNA, a *p53* mutation in a 1.7 kb amplicon, and unknown *p53* mutations in pooled DNA samples. EndoV/Ligase mutation scanning combined with PCR/LDR/Universal array proved superior to automated DNA sequencing for detecting *p53* mutations in colon tumors. This technique is well suited for scanning low-frequency mutations in pooled samples and for analysing tumor DNA containing a minority of the unknown mutation.
Oncogene (2002) **21,** 1909–1921. DOI: 10.1038/sj/onc/1205109

Keywords: mutation detection; endonuclease V; thermostable ligase; mismatch recognition

Introduction

Cancers arise from the accumulation of inherited and sporadic mutations in cell cycle, DNA repair, and growth signaling genes. Knowledge of the inherited or sporadic change in tumor suppressor genes or oncogenes could influence prevention and treatment decisions. For example, certain specific single nucleotide polymorphisms (SNPs) in cancer genes such as *BRCA1, BRCA2* or *APC* are present at low frequency (1–5%) in certain ethnic groups. These carriers are at significantly higher risk of developing breast, ovarian, prostate or colon cancers (Abeliovich *et al*., 1997; Beller *et al*., 1997; Berman *et al*., 1996; Laken *et al*., 1997; Oddoux *et al*., 1996; Roa *et al*., 1996; Struewing *et al*., 1995, 1997). Similarly, those with inherited mutations in the von Hippel–Lindau gene (*VHL*) would benefit from early screening and preventive surgery at the appropriate time (Chen *et al*., 1996; Stolle *et al*., 1998; Libutti *et al*., 2000; Mohr *et al*., 2000; Sgambati *et al*., 2000; Walther *et al*., 1999). Sporadic mutations in the *p53* gene influence both clinical outcome and response to therapy (Broll *et al*., 1999; Bunz *et al*., 1999; Dameron *et al*., 1994; Heide *et al*., 1997; Prives and Hall, 1999; Tortola *et al*., 1999; Zou *et al*., 2000). The precise nature of the p53 mutation presents both challenges and opportunities for alternate treatment strategies in specific cancers (Wang *et al*., 1998b; Aurelio *et al*., 1998, 2000; Foster *et al*., 1999; Webley *et al*., 2000). These diverse genes highlight the clinical need to accurately identify often unknown inherited aberrations or infrequently represented mutations in mixed populations of DNA molecules.

A variety of methods have been developed to successfully scan for unknown mutations. Unfortunately, current technologies used for *in vitro* scanning are typically lacking in either throughput capacity or sensitivity to rapidly detect a range of mutations. This can result in false-negatives and lower throughput, since many of these techniques cannot detect tumor mutations in the presence of contaminating stromal cells, or be amenable to the use of pooled samples. Cycle-sequencing can detect any mutation and its position (Innis *et al*., 1988; Tabor and Richardson, 1995), but this method has low sensitivity which prevents an accurate analysis of pooled samples. Automated fluorescent dideoxysequencing often failed to detect germline mutations in DNA repair genes (Yan *et al*., 2000a,b). Variation detection arrays (VDA)

*Correspondence: F Barany; E-mail: barany@med.cornell.edu
[4]Current address: Department of Genetics and Biochemistry, South Carolina Experiment Station, Clemson University, 122 Long Hall, Clemson, SC 29634-0324, USA.
Received 13 August 2001; revised 15 October 2001; accepted 29 October 2001

use standard hybridization microarrays to scan large sequence blocks in given genes, however, false positive rates of 11–21% have been observed, and there are particular difficulties in detecting frameshift mutations (Cargill *et al*., 1999; Halushka *et al*., 1999; Wang *et al*., 1998a; Hacia *et al*., 1996; Hacia, 1999). A direct comparison between commercially available *p53* hybridization chips and automated fluorescent dideoxysequencing showed they miss 23% and 25% of mutations in tumors, respectively (Ahrendt *et al*., 1999). These techniques have reached their theoretical limits due to the dilution of mutant DNA in tumor samples by stromal cell infiltration.

Unknown mutations may be detected by resolving homoduplex and heteroduplex DNA based on their differing electrophoretic migration using single-stranded conformational polymorphism (SSCP) (Hayashi, 1991; Korn *et al*., 1993; Makino *et al*., 1992; Suzuki *et al*., 1990), denaturing-gradient gel electrophoresis (DGGE) (Fahy *et al*., 1997; Fodde and Losekoot, 1994; Guldberg and Guttler, 1994; Ridanpaa and Husgafvel-Pursiainen, 1993; Ridanpaa *et al*., 1995), constant denaturing capillary electrophoresis (CDCE) (Chen and Thilly, 1994; Khrapko *et al*., 1994), dideoxy fingerprinting (ddF) (Sarkar *et al*., 1992), and restriction endonuclease fingerprinting (REF) (Liu and Sommer, 1995), and more recently, denaturing high-performance liquid chromatography (DHPLC) (Underhill *et al*., 1996). Each of these approaches has their advantages, nevertheless, none of the approaches has the combined ability to rapidly detect low level mutations, distinguish missense from silent polymorphisms, and locate the position of the mutation. Polymorphisms may also be identified by cleavage of mismatches in DNA-RNA hybrids via RNase A mismatch cleavage (Myers *et al*., 1985; Perucho *et al*., 1989; Winter *et al*., 1985), as well as in DNA-DNA homoduplexes via chemical cleavage (Cotton *et al*., 1988; Hansen *et al*., 1995; Haris *et al*., 1994), or enzymatic cleavage by T4 Endonuclease VII, CEL1, or MutY (Giunta *et al*., 1996; Oleykowski *et al*., 1998; Xu *et al*., 1996; Youil *et al*., 1995). These enzymatic cleavage approaches identify the approximate position of most polymorphisms. However, these enzymes often also nick matched DNA leading to high backgrounds and limiting their usefulness in identifying DNA lesions in solid tumors. In an ideal mutation scanning method, no or little cleavage should be observed in a completely matched DNA, while a wide variety of mismatches should result in cleavage (Taylor, 1999).

To overcome background signal from enzymatic cleavage, we developed a mutation scanning method which combines the mismatch recognition and nicking ability of thermostable Endonuclease V (EndoV) with the high fidelity of a thermostable DNA ligase to proofread and reseal background nicks at duplex DNA (see Figure 1). Endonuclease V recognizes a wide variety of DNA lesions, predominantly deoxyinosine, xanosine, uracil, and apurinic sites, and under select buffer conditions, base mismatches and insertions/deletions (Huang *et al*., 2001; Yao and Kow, 1994; Liu *et al*., 2000). The enzyme nicks DNA one base 3′ to the mismatch in heteroduplexed DNA (Figure 1). As with all mismatch recognition enzymes, EndoV will also nick matched regions of the DNA, but in contrast to other enzymes, EndoV leaves a ligatable end with a 5′ phosphate (Yao *et al*., 1994; Yao and Kow, 1994). We have cloned a thermostable ligase from *Thermus* species AK16D that exhibits 2.5 to 5-fold better discrimination at the 3′ penultimate position compared to other ligases (Tong *et al*., 1999). By using these two enzymes sequentially, we can religate spurious nicks, while maintaining the desired nicks at the polymorphic sites, resulting in a highly sensitive mutation scanning assay. In this report, we describe an optimized scheme for detection of cancer-associated mutations using *Tma*. EndoV in conjunction with *Tsp*. AK16D Ligase. PCR fragments up to 1.7 kb were tested and correct mutations can be detected in the entire fragment size range. The approximate position of the mutation is determined from the fragment length, and confirmed by sequencing. The assay can detect 98% of the typical mutations found in the human genome, with only two GC-rich sequences refractory to detection. It is also highly sensitive and can identify a mutant sequence at a 1 : 20 dilution in wild-type DNA, making it amenable to pooling strategies.

Results

Strategy, optimization and characterization of EndoV/ligase mutation scanning assay

Initial characterization of *Tma* EndoV activity was performed on fluorescently labeled synthetic substrates containing mismatched bases. Optimal cleavage at the mismatch with minimal non-specific cleavage was achieved in a 10 mM HEPES (pH 7.4) buffer containing, 1 mM DTT, 2% glycerol, and 5 mM $MgCl_2$ (Huang *et al*., 2001). Analysis of cleavage patterns of all 12 natural mismatched base-pairs suggests that purine bases are preferentially cleaved, showing a general hierarchy of A = G > T > C (Table 1). This hierarchy is based on the amount of cleavage product observed for each strand of the mismatched heteroduplexed substrates.

To test the utility of the EndoV/ligase scheme for finding mutations in the genome, cell line DNA containing either wild-type, G12V, or G12D mutations in the K-*ras* gene were PCR amplified (see Table 2 for list of PCR primers used). Mutant amplicons were mixed with wild-type fragments, denatured, reannealed, and treated with EndoV. The single nucleotide change G12V generates G/A and T/C mismatches, while G12D generates G/T and A/C mismatches. Initial results in the above buffer yielded weak bands, however, cleavage of the heteroduplex DNA was enhanced by the addition of 5% DMSO and 1–1.5 M N,N,N-trimethylglycine (betaine, data not shown). The EndoV cleaved both strands for G12V/G12wt and G12D/G12wt heteroduplexes, resulting in products of approximately 159 nt and 116 nt (Figure 2). Without the addition of DNA ligase, a faint, larger, non-specific cleavage fragment was observed (Figure

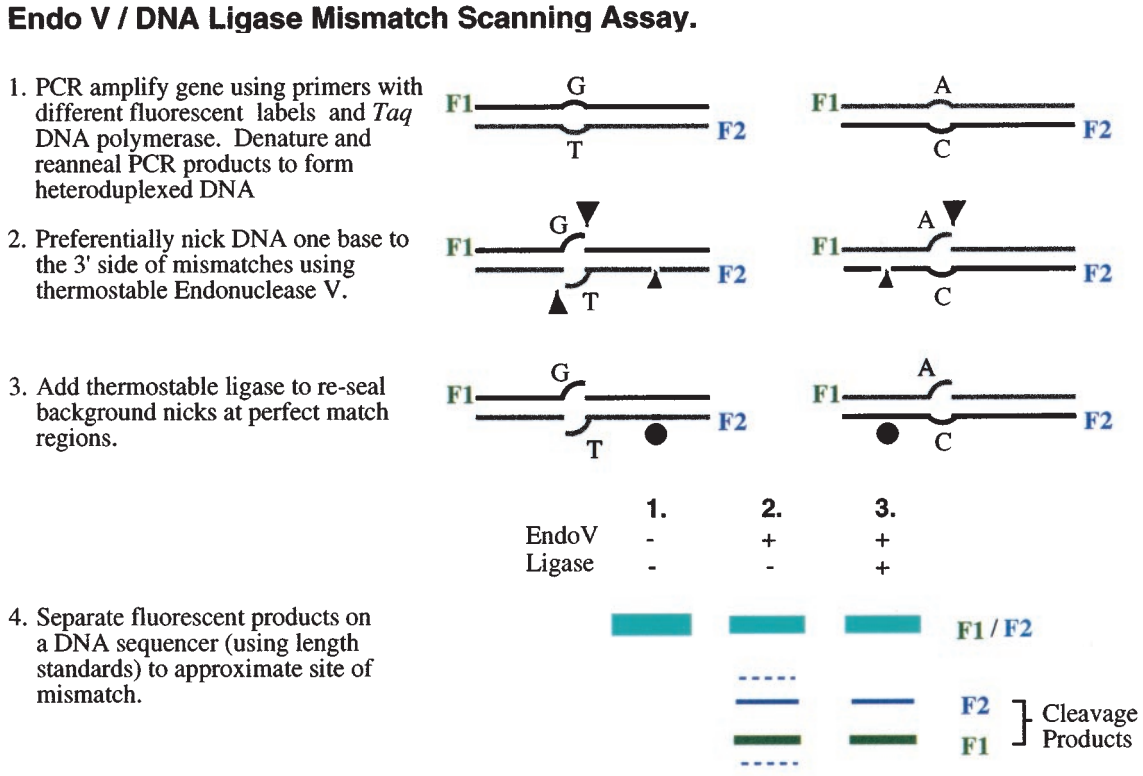

Figure 1 Endonuclease V/DNA ligase mismatch scanning assay for scoring unknown mutations. Heteroduplexes are formed from PCR amplicons containing both normal and variant sequence. Mixed amplicons of variant sequence can be obtained by amplifying heterozygous germline samples, or from tumor samples where stromal cell contamination provides sufficient amounts of wild-type DNA, or by mixing the PCR products from unknown and wild-type samples in a 1 : 1 ratio. *Tma* EndoV nicks DNA one base 3′ to the mismatch (big triangle), and it can also generate non-specific nicks in homoduplex DNA with minor activity (small triangle). Addition of *Thermus sp*. AK16D DNA ligase (solid circle) seals these non-specific nicks, providing a proofreading mechanism to improve signal-to-noise. Both top and bottom strand PCR primers are 5′ end-labeled with different fluorescent dyes (6-FAM and TET, respectively) allowing for cleavage products to be distinguished on a denaturing polyacrylamide gel. The approximate position of the mutation can be determined from the resultant fragment lengths

Table 1 Summary of *Tma* endonclease V cleavage of heteroduplexed synthetic single-base mismatched substrates[a]

Base change (*Wt*←→*Mt*)	*A G* |←→| *T C*	*C T* |←→| *G A*	*A C* |←→| *T G*	*G T* |←→| *C A*	*A T* |←→| *T A*	*G C* |←→| *C G*
Hetroduplex I:						
Upper Strand (Wt)	A + + +[b]	C +	A + +	G + +	A + + +	G+ + +
Bottom Strand (Mt)	C −	A+ +	G + + +	A + +	A + + +	G + + +
Heteroduplex II:						
Upper Strand (Mt)	G + +	T + +	C +	T + +	T +	C −
Bottom Strand (Wt)	T + +	G + + +	T + +	C −	T +	C −

[a]Upper Strand:
5′–Fam–TACCCCAGCGTCTGCGGTGTTGCGTNAGTTGTCATAGTTTGATCCTCTAGTCTTGTTGCGGGTTCC–3′
3′–GGGGTCGCAGACGCCACAACGCANTCAACAGTATCAAACTAGGAGATCAGAACAACGCCC–Tet–5′
Bottom Strand:
[b]Cleavage symbols: (+ + +)=about 10%; (+ +)=about 5%; (+)=about 2%; (−) no cleavage observed

2). However, this fragment was significantly reduced when the cleavage products were incubated with *Tsp*. AK16D ligase (compare first four lanes with last 12 lanes). Thus, non-specific cleavage nicks can be preferentially sealed by the DNA ligase, which acts as a proofreading enzyme in this assay.

Some low molecular weight non-specific fragments still persist following the ligation step (bottom, Figure 2). They could be reduced by filtration of the PCR sample prior to heteroduplex formation using Microcon 30 spin columns (data not shown). Filtration removes most unused primers, which can hybridize to

Table 2 PCR primers to amplify exons of the K-*ras*, *APC*, *p53*, *BRCA1*, *BRCA2*, and *VHL* genes

Gene	*Exon*	*Strand*	*Primer sequence*
K-*ras*	Exon 1	Top	Tet- 5′-CCCCATAGTGTATTAACCTTATGTGTGACATGTTC-3′
		Bottom	Fam- 5′-CCCCAAAATGGTCAGAGAAACCTTTATCTGTATC-3′
APC	Exon 15	Top	Tet-5′-CCCCGCTGCCACTTGCAAAGTTTCTTC-3′
		Bottom	Fam-5′-CCCCACTCTGAACGGAGCTGGCAAT-3′
p53	Exon 5	Top	Tet-5′-CCCCTGTTCACTTGTGCCCTGACTTTC-3′
		Bottom	Fam-5′-CCCCCAGCTGCTCACCATCGCTATC-3′
	Exon 6	Top	Tet-5′-CCCCTCTGATTCCTCACTGATTGCTCTTA-3
		Bottom	Fam-5′-CCCGGCCACTGACAACCACCCTTAAC-3′
	Exon 7	Top	Tet-5′-CCCGCCTCATCTTGGGCCTGTGTTATC-3′
		Bottom	Fam-5′-CCCGTGGATGGGTAGTAGTATGGAAGAAAT-3′
	Exon 8	Top	Tet-5′-CCCGGACAGGTAGGACCTGATTTCCTTAC-3′
		Bottom	Fam-5′-CCCCGCTTCTTGTCCTGCTTGCTTAC-3′
	1.7kb	Top	Fam-5′-CCCGCATGGTGGTGCACACCTATAGTC-3′
		Bottom	Tet-5′-CCCAAGCTGTTCCGTCCCAGTAGATTAC-3′
BRCA1	Exon 2	Top	Tet-5′-CCCCTCATTGGAACAGAAAGAAATGGATTTATC-3′
		Bottom	Fam-5′-CCCCTCTTCCCTAGTATGTAAGGTCAATTCTGTTC-3′
	Exon 20	Top	Tet-5′-CCCCACTTCCATTGAAGGAAGCTTCTCTTTC-3′
		Bottom	Fam-5′-CCCCATCTCTGCAAAGGGGAGTGGAATAC-3′
BRCA2	Exon 11	Top	Tet-5′-CCCCCAAAATATGTCTGGATTGGAGAAAGTTTC-3′
		Bottom	Fam-5′-CCCCTTGGAAAAGACTTGCTTGGTACTATCTTC-3′
VHL	Exon 1	Top	Tet-5′-CCCGACCGCGCGCGAAGACTAC-3′
		Bottom	Fam-5′-CCCAGGGGCTTCAGACCGTGCTATC-3′
	Exon 2	Top	Tet-5′-CCCCACCGGTGTGGCTCTTTAACAAC-3′
		Bottom	Fam-5′-CCCAGGGGCTTCAGACCGTGCTATC-3′
	Exon 3	Top	Tet-5′-CCCTAGTTGTTGGCAAAGCCTCTTGTTC-3′
		Bottom	Fam-5′-CCCAAACTAAGGAAGGAACCAGTCCTGTATC-3′

the ends of single-stranded PCR products during the heteroduplex formation step. EndoV cleaves at the single-strand region of these PCR fragments to generate low molecular weight non-specific bands. This artifact can be alleviated by either removing unused primers, or designing the primers 30 bp outside the anticipated region of mutation scanning, such that lower molecular weight fragments do not interfere with interpretation of the results.

Most of the current gel-based mutation scanning methods are limited to short PCR fragments. To assess the application of the EndoV/ligase assay on long PCR fragments, 1.7 kb fragments containing either wild-type or *p53* R248W mutations were amplified. Fragments containing this mutation generated a unique 413 nt cleavage product that was not observed in the pure wild-type sample (Figure 3). In addition, treatment of the cleavage products with ligase resulted in reduction of most non-specific cleavage bands. Remaining bands (i.e. in lanes 5, 7 and 8) represent putative SNPs, but this needs to be confirmed by sequencing. As is often the case with C→T transitions, cleavage of only one strand (containing the G→A change) by EndoV was observed.

To determine the sensitivity of the EndoV/ligase assay, PCR fragments containing K-*ras* G12V, G12D, and G12A mutations were diluted with a wild-type PCR fragment in ratios ranging from 1 : 1 to 1 : 100. The relative fluorescence intensity of EndoV cleavage products as a function of mutant-to-wild-type fragment ratios were determined (Figure 4). Using the criterion that cleavage signals twofold above background are significant, these results indicate that the cleavage signal can be detected in samples diluted 20-fold with wild-type PCR fragments. This result suggests that this assay is compatible with sample pooling.

Detection of various point mutations in p53 and VHL genes

To study the versatility of *Tma* EndoV in scanning various point mutations, we applied this technique to both tumor and germline samples containing different known point mutations in *p53* and *VHL* (Figure 5 and Table 3). These genes and exons were chosen because mutant sequences are difficult to detect (Sgambati *et al*., 2000; Ahrendt *et al*., 1999). Addition of ligase significantly reduced non-specific cleavage background (compare *p53* mutation detection in Figure 5, lanes 13–18 with 19–24). EndoV/Ligase mutation scanning detected mutations in 9/11 of the *p53* samples and 12/13 of the *VHL* samples. The refractory sequences were transition mutations in GC-rich DNA (g<u>R</u>cg and rc<u>R</u>c, where R = purine, and the underlined base is the position of the mutation). The 500 bp *VHL* exon 1 fragment did not amplify efficiently with the standard PCR methodology, possibly due to the very high GC content in the middle of this exon. Therefore, we diluted the PCR products 100-fold and used 1 μl of this diluted reaction as template for a second round of amplification under the same PCR conditions, but for only 15 cycles. This resulted in efficient amplification and detection of the point mutations. Overall, with the exception of a few refractory sequences, the scanning assay detected point mutations in GC-rich DNA.

Detection of the T→A mutation in codon 1307 of the APC gene

To validate the utility of this assay on mononucleotide repeats, we applied this method to germline samples scanning the *APC* I1307K mutation, which creates an

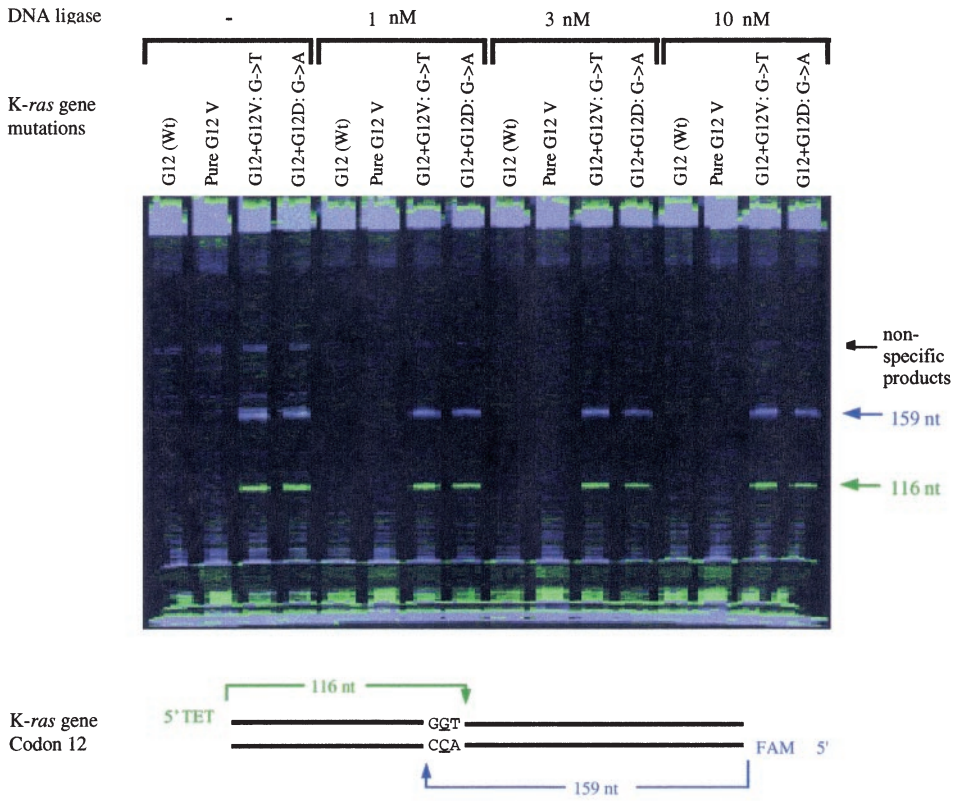

Figure 2 Detection of mutations in K-*ras* gene codon 12 using *Tma* EndoV and *Tsp*. AK 16D DNA ligase. PCR fragments of K-*ras* exon 1 from genomic DNA of pure wild-type (G12) and pure mutant (G12V) were used as controls. The mixtures of PCR fragments of wild-type and pure mutant (G12V)(1 : 1 ratio) or wild-type and pure mutant (G12D)(1 : 1 ratio) were used as DNA substrate samples in the *Tma* EndoV cleavage assay. DNA duplexes (100 ng) were cleaved in a buffer containing 500 nM *Tma* EndoV and other reagents as described in Materials and methods. Increasing amounts of *Tsp*. AK16D DNA ligase were added to seal the nicks generated by non-specific cleavage, as described in Materials and methods. The ligation mixtures were electrophoresed in a 6% denaturing polyacrylamide gel. Top and bottom strand PCR primers were 5′ end-labeled with TET and 6-FAM (showing green and blue, respectively, using filter C setting on an ABI 377 DNA sequencer). The green band was generated from cleavage of the top strand, and the blue band was generated from the cleavage of the bottom strand

A_8 mononucleotide repeat (Laken *et al*., 1997). The DNA mutations in these samples were previously verified by PCR/LDR (Zirvi *et al*., 1999). Nine samples from patients carrying the mutation and four from normal individuals were assayed. Mutations in samples from all of the carriers were detected, with no false positives detected in the normal samples (data not shown). It was found that cleavage activity is higher in the presence of 1.0 M betaine instead of 1.5 M betaine (data not shown). This is presumably due to the higher AT content of exon 15 in *APC* than previous genes we assayed (see Discussion). Therefore, the optimum amount of betaine added to a reaction mixture may vary depending on the AT content of the PCR fragment being amplified.

Detection of small deletion and insertion mutations in BRCA1, BRCA2 and VHL genes

In order to study the utility of this assay to detect small deletions or insertions, segments of *BRCA1* and *BRCA2*, *VHL* genes containing 1–3 nt insertions/deletions were amplified. Cleavage of single-base (five cases including *p53* samples in Table 4 below), two-base (two cases) or three-base (one case) insertions or deletions generally gave robust signal, with the exception of the single-base insertion and deletion found in *BRCA1* and *BRCA2* (Figure 5 and Table 3). In the case of *BRCA2* 6174delT, addition of ligase eliminates a background band that is stronger than the correct cleavage product at the mismatch (compare lanes 9 and 10, with the incorrect band at 140 bp with lanes 11 and 12 showing only the correct band at 151 bp). For the *BRCA1* 185delAG mutation, multiple strong bands on both strands are observed presumably reflecting alternate structures of the heteroduplexed substrates (lanes 2 and 4). Thus, insertions and deletions that commonly create frameshift mutations are distinguished by the EndoV/ligase scanning assay.

Detection of p53 mutations in colon tumor samples

An analysis of *p53* mutations in DNA isolated from 50 Stage I–IV colon adenocarcinomas using a combination of both PCR/LDR/Universal array and EndoV/Ligase mutation scanning proved more accurate than automated sequencing alone (see Table 4, Favis *et al*., unpublished work; Favis *et al*., 2000; Gerry *et al*.,

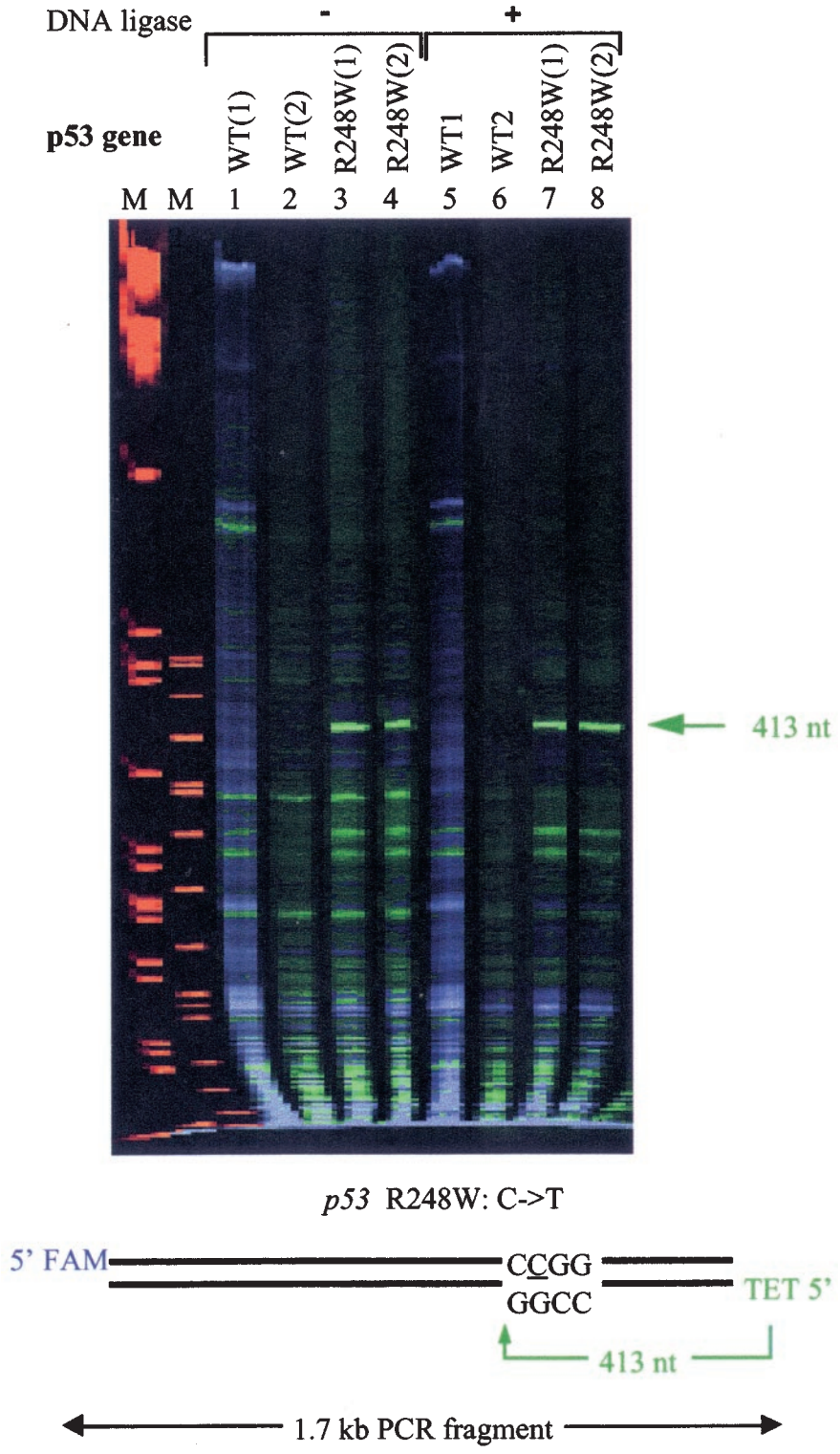

Figure 3 Using *Tma* EndoV and *Tsp.* AK16D DNA ligase to detect the *p53* R248 C→T mutation in a 1775 bp PCR fragment. Wild-type PCR fragments from two different individuals, WT1 and WT2, were used as normal controls. Mutant PCR fragments containing R248 C→T mutations from two different tumors were used as samples in the mutation scanning assay. The reaction mixtures were electrophoresed in a 6% denaturing polyacrylamide gel. Lanes 1 to 4 without ligase, and lanes 5 to 8, with 6 nM *Tsp.* AK16D DNA ligase. M1: GeneScan 2500 molecular standard, M2: GeneScan 500 molecular standard

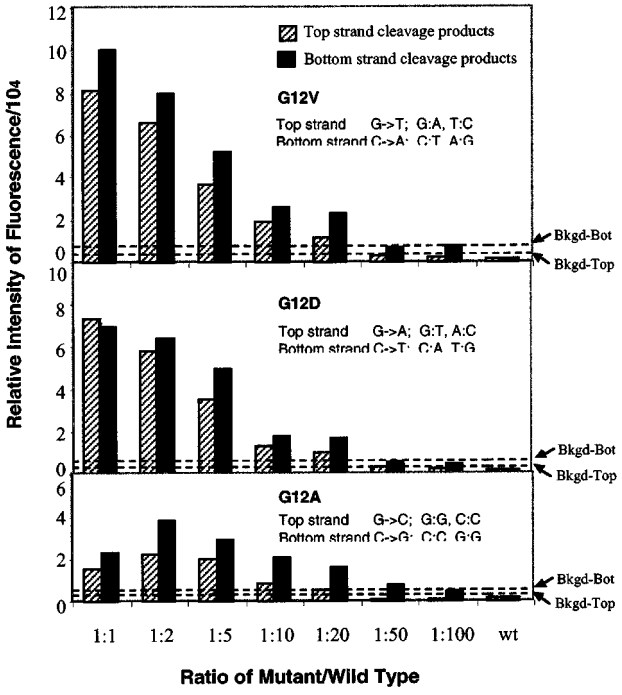

Figure 4 Sensitivity of the Endo V/DNA ligase reaction. Wild-type and mutant K-*ras* exon 1 was PCR amplified using genomic DNA from the following cell lines: SW480 (pure G12V mutant), LS180 (1.8 pure G12D mutant : 1 wild-type), SW1116 (0.7 pure G12A mutant : 1 wild-type) and HT29 (pure wild-type). Mutant amplicons were diluted with wild-type amplicons to produce mutant : wild-type ratios ranging from 1 : 1 to 1 : 100 (note that for G12A, the actual value of the first ratio in the graph is 0.7 : 1). The heteroduplex PCR fragments were cleaved by *Tma* EndoV and ligated with *Tsp*. AK16D. The signals that migrated similarly to cleavage products from the wild-type homoduplex were used as wild-type (wt) controls. Two background peaks located near either the top or bottom strand cleavage products were used as background signals to compare signal and noise for the appropriate cleavage product. The average background signal (represented by dashed lines) was calculated as the mean value of background signals from seven samples with different mutant : wild-type ratios and one wild-type sample. Bar graphs indicate the relative fluorescence intensity of top strand (striped bars) and bottom strand (solid bars) cleavage products for each mutant : wild-type ratio. The relative fluorescence intensity is defined as the area under a signal's peak as determined by GeneScan analysis software. Bkgd-top: average background signals for the top strand cleavage products; bkgd-bot: average background signals for the bottom strand cleavage products. The mutation, nucleotide change, and the mismatched base pairs are indicated for each panel. The position of each mutation is underlined in the mismatch base pairing

1999). Of the 23 samples, 26 *p53* mutations were identified by PCR/LDR combined with EndoV/ligase, while eight samples were missed by automated sequencing (65% true positive, 35% false negative). Confirmation of the exact sequence in the samples called positive with EndoV/Ligase mutation scanning, required gel purification of some PCR fragments, and re-sequencing of both strands with manual reading. Significantly, EndoV/Ligase mutation scanning scored all four frameshift mutations, which accounts for almost 20% of the samples with *p53* mutations. Such frameshift mutations are beyond the detection capacity of commercially available *p53* hybridization chips (Ahrendt *et al*., 1999).

The above samples were pooled in subsets to test the ability of the assay in a mixed sample format. These tumor samples were frozen directly in liquid nitrogen and tissue sections were not microdissected prior to extraction of nucleic acids. Microscopic examination of sections revealed stromal cell infiltration ranging from about 10–50%. Mutation of one *p53* allele is invariably accompanied by loss of heterozygosity (LOH) of the other allele, occurring either through partial to complete chromosome loss or through mitotic non-disjunction (i.e. both mutant chromosomes migrate to one daughter cell) (Thiagalingam *et al*., 2001). Thus, DNA from the above samples would have a range of mutant p53 : wild-type allele of from about 90 : 10% (for non-disjunction with 10% stroma) to 33 : 67% (for chromosome loss with 50% stroma). When the above samples were re-analysed for *p53* mutations in pools of three samples together, the EndoV/ligase mutation scanning assay could still distinguish the presence of all the mutants. Several of the pooled sample bands were even stronger than from individual samples (suggesting the original sample was mostly mutant DNA), while a few bands were weaker but could still be detected (suggesting the original sample had substantial stromal contamination). The ability to detect mutations in pools of five or even 10 samples demonstrates the sensitivity of the assay. This high sensitivity is potentially useful in routine analysis of individual clinical samples in which relatively few neoplastic cells are present.

Discussion

Mutation scanning requires robust techniques with high specificity and sensitivity. Current enzymatic methods often suffer from low specificity or low sensitivity (Taylor, 1999). Endonuclease V exhibits broad mismatch cleavage activities and generates a ligatable nick at the 3′ penultimate position (Yao and Kow, 1994). DNA ligases vary substantially as to their mismatch ligation fidelity, but a thermostable ligase from *Thermus* species AK16D demonstrates superior fidelity in sealing a matched nick at the 3′ penultimate position (Tong *et al*., 1999). By combining the cleavage/ligation activities of these two enzymes, we developed a mutation scanning method where ligase proofreading reduced background and increased sensitivity.

The utility and versatility of this assay has been demonstrated on over 80 samples containing known and unknown SNPs, or mutations in *p53, VHL*, K-*ras, APC, BRCA1*, and *BRCA2* (see Tables 3 and 4). Most point mutations (31/35 unique mutations) and all frameshift mutations (8/8 unique insertions or deletions) could be detected with this method (Figures 2 and 5, Tables 3 and 4). The four refractory mutations, K-*ras* G13D(G→A), *p53* R175H(G→A), R273C(C→T) and *VHL* G164D(G→A) are all transition mutations of the

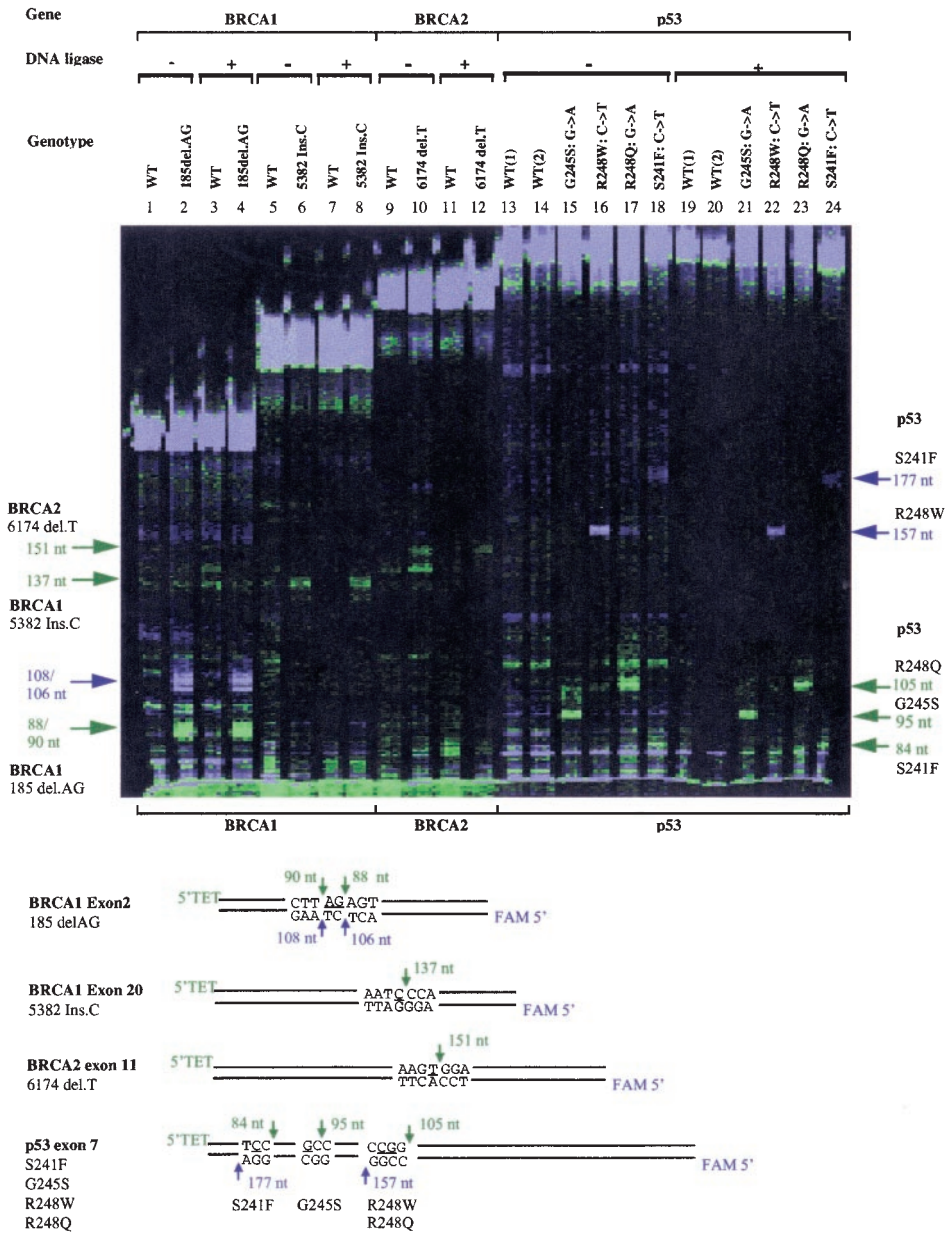

Figure 5 Detection of small insertions/deletions in *BRCA1, BRCA2* and point mutations in *p53*. Lanes 1–12 show Endo V/Ligase reactions that detect the three founder mutations in *BRCA1* and *BRCA2*. Lanes 13–24 show assays for mutations in exon 7 of *p53*. The results of incubation with and without DNA ligase are shown. Mutation tested and presence or absence of ligase are indicated above gel lanes. To estimate site of cleavage, forward and reverse PCR primers were 5′ end labeled with TET and 6-FAM, respectively. Multiple cleavage bands with 185 delAG mutation reflects different conformations of heteroduplexed DNA. Arrows indicate positions of bands of interest

form gRcg and rcRc (Table 3). These refractory polymorphisms appear on average only 2% of the time in human DNA, as judged by their frequency in 6000 random SNPs culled from the public SNP database. Further, combining PCR/LDR detection of the known refractory sequences with EndoV/Ligase mutation scanning can provide essentially complete coverage of all point and frameshift mutations in a given gene (Table 4).

For detectable mutations resulting from a transition, the purine strand at the mismatch is always cleaved, while the pyrimidine strand is generally not cleaved. Of the 22 transition mutations assayed, in only six cases were both strands cleaved (Table 3). Among the seven transversions assayed, both the top and bottom strands containing both purine and pyrimidine bases were cleaved for all but one case studied. (The top strand of F190L produced no observable product exactly 56 bp in length, but there were some non-specific bands around this position.) We did not distinguish between purine and pyrimidine strand cleavage for transversions. These results follow the general hierarchy for *Tma*

Table 3 Summary of *Tma* EndoV/Ligase mutation scanning of cancer genes

Gene	*Exon*	*Mutation*	*Surrounding sequences*	*Change*	*Cleavage*[a] *Top strand*	*Bottom strand*
K-*ras*	Exon 1	G12V	TGGTG	G→T	+ + +	+ +
		G12D	TGGTG	G→A	+ +	+ +
		G12A	TGGTG	G→C	+ +	+
		G13D	TGGCG	G→A	−	−
APC	Exon 15	I1307k	AATAA	T→A	+	+ +
p53	Exon 5	C135Y	TTGCC	G→A	+ +	−
		R175H	GCGCT	G→A	−	−
	Exon 6	R196C	TCCGA	C→T	−	+ +
		Y220C	CTATG	A→G	+ +	−
	Exon 7	S241F	TTCCT	C→T	+	+
		G245S	GCGGC	G→A	+ + +	−
		R248Q	CCGGA	G→A	+ + +	−
		R248W	ACCGG	C→T	−	+ +
	Exon 8	R273H	GCGTG	G→A	+	+
		R273C	TGCGT	C→T	−	−
		R282W	ACCGG	C→T	+ +	+ +
BRCA1	Exon 2	185 del.AG	TT*AG*AG	AG deletion	+ + +	+ + +
	Exon 20	5382 Ins.C	AT*C*CC	C insertion	+	−
BRCA2	Exon 11	6174del.T	AG*T*GG	T deletion	+	−
VHL	Exon 1	P157L	GCCCG	C→T	−	+ +
		W159A	TATGG	T→A	+ + +	+ +
		G164D	CGGCG	G→A	−	−
		Y169H	CCTAC	T→C	−	+
		Y183H	GCTAC	T→C	+ +	+ +
	Exon 2	F190L	TTCAG	C→G	−	+ +
		G198G	GGGCT	G→T	+ +	+ + +
		L199F	GGCTT	C→T	−	+ + +
		S200 Ins. E	CT*AGA*GG	AGA insertion	+ + +	+ + +
		A220T	TTGCC	G→A	+ + +	−
		N211 del. A	CC*A*AT	A deletion	+ + +	+ + +
	Exon 3	L229P	TCTGA	T→C	−	+
		R232Q	GCGAT	G→A	+	+ +
		R238W	TCCGG	C→T	−	+ +
		L259Q	TCTGG	T→A	+ + +	+ +

[a]Cleavage symbols: (+ + +) = about 10%; (+ +) = about 5%; (+) = about 2%; (−) no cleavage observed

EndoV cleavage when using synthetic oligonucleotides as substrates (Table 1, Huang *et al*., 2001). The intensity of cleavage of one strand follows A = G > T > C, in which G, A, T and C indicates the base in the mismatch (Tables 1 and 3). *Tma* EndoV cleavage of the strand containing C in the mismatch was usually very poor and sometimes not observed at all.

The cleavage activity of *Escherichia coli* (*E.coli*) EndoV is affected by the sequence content surrounding the mismatch. The presence of two G/C pairs, one 5′ and the other 3′ immediately adjacent to the mismatch abolishes the ability of the enzyme to cleave the mismatch (Yao and Kow, 1994). It appears that *Tma* EndoV cleavage activity is not significantly constrained by this neighbor effect. We assayed 14 mutations flanked with two G/C pairs, one 5′ and the other 3′ to the mismatch. Of these, *Tma* EndoV cleaved 10 mutations, often giving robust signal. Besides the structural differences between the two EndoV proteins, assaying at higher temperature (65°C) and the addition of betaine or DMSO to the reaction buffer may have facilitated *Tma* EndoV cleavage of DNA containing mismatches flanked by two G/C pairs. Betaine has been used to equalize the melting temperatures of DNA fragments having different GC content (Rees *et al*., 1993), as well as an additive to facilitate the PCR amplification of DNA regions with high GC content (Henke *et al*., 1997). Knowledge of the AT content of the fragment should guide selection of a betaine concentration that provides optimal cleavage while minimizing background.

A two or three base deletion/insertion resulted in cleavage products with very strong signal for both strands (*BRCA1* 185 delAG and *VHL* S200 insAGA), and a strong single band in one case (*p53* Q167 delGT). An A deletion in *VHL* also generated a strong signal for both strands. An A or T deletion in *p53* generated a moderate signal for one strand, while a C insertion in *BRCA1* and a T deletion in *BRCA2* only resulted in a weak signal for one strand, and no signal for the other. These results suggest that hybrids with multiple base insertions/deletions are good substrates for EndoV, whereas one base insertions or deletions are less efficiently cleaved and surrounding sequence may have an effect. Nevertheless, EndoV/ligase scanning identified all the frameshift mutations tested.

With this method we can detect a mutation in PCR fragments as long as 1.7 kb. In addition to the known mutation, multiple cleavage bands were observed even after the ligation step. Since these additional bands are also present in one but not the other wild-type DNA samples (WT1, Figure 3), this result is consistent with

Table 4 Detecting *p53* mutations in tumor samples: Comparisons of using a combined analysis of PCR/LDR and Endo V/ligase mutation scanning to Dideoxysequencing

Sample #: (50 total)	*PCR/LDR Universal array*	*Endo V/Ligase Mutation scanning*	*DNA sequencing, Automated read*	*Resequencing of booth strands, manual read*
53	(5)[a] R175 G2-A		(5) R175 G2-A	
55	(8) E285 G1-A	(8) E285 G1-A	(8) E285 G1-A	
59	(8) R273 C-T		(8) R273 C-T	
60	(8) R282 C-T	(5) K174 A-T		(5) K174 A-T
	(8) R306 C-T			(8) R306 C-T
65	(5) R175 G2-A		(5) R175 G2-A	
66	(7) R248 G-A	(7) R248 G-A		(7) R248 G-A
67	(8) R282 C-T	(8) R282 C-T		(8) R282 C-T
68		(6) Y205 A-T		(6) Y205 A-T[b]
71	(8) R273 C-T		(8) R273 C-T	
73	(8) R282 C-T	(5) Q167 delGT,	(8) R282 C-T	Q167 delGT[b]
		(8) R282 C-T		(8) R282 C-T
77		(7) S261 T-G	(7) S261 T-G	
78	(7) R248 G-A	(7) R248 G-A		(7) R248 G-A
79		(7) S240 delA		(7) S240 delA[b]
80	(5) R175 G2-A		(5) R175 G2-A	
81	(5) R175 G2-A		(5) R175 G2-A	
84		(7) S261 T-G	(7) S261 T-G	
89		(6) Q192 C-T	(6) Q192 C-T	
90		(6) H214 delT		(6) His 214 delT[b]
93	(5) R175 G2-A		(5) R175 G2-A	
94		(6) H214 delT		(6) His 214 delT[b]
96	(8) R273 G-A	(8) R273 G-A	(8) R273 G-A	
97		(8) R276 C-G	(8) R276 C-G	
98	(8) R282 C-T	(8) R282 C-T	(8) R282 C-T	
(27 samples)	No Mutation	No Mutation	Not determined[c]	Not determined[c]
Score	15/23	16/23	15/26	20/26
	(16 known, 4 new, 4 deletions)	(2 resistant sites, observed 6 times)		
Per cent	65%	70%	57%	77%
Combined	100%		83%	
Adjusted	16/16	16/17	15/23	23/23
Per cent[d]	100%	94%	65%	(100%)

[a]The numbers in parenthesis refer to the *p53* exon where the mutations are located. [b] Required gel purification of PCR product to obtain sequencing result. [c] Sequencing of five random samples of 27 called negative by PCR/LDR and Endo V/Ligase reveal no new mutations. [d] Adjusted Per cent reflects the ratio of mutations detected/mutations the technique is designed to detect

additional polymorphic sites being present in the tumor sample. These combined results indicate that this assay can detect a mutation in fragments up to 1.7 kb in length, and suggests that multiple mutations can be identified in a single amplicon. In addition, dilution experiments with K-*ras* and pooling experiments with *p53* from tumor samples demonstrate that this assay is amendable to pooling. Therefore, the throughput of EndoV/ligase mutation scanning has the potential of being increased by both scanning larger regions of DNA in a single reaction as well as scanning multiple samples within the same reaction.

Tests for germline *BRCA1, BRCA2, APC*, and *VHL* mutations provide highly accurate disease diagnosis or prediction of future disease (Abeliovich *et al*., 1997; Beller *et al*., 1997; Berman *et al*., 1996; Laken *et al*., 1997; Oddoux *et al*., 1996; Roa *et al*., 1996; Struewing *et al*., 1995, 1997; Chen *et al*., 1996; Stolle *et al*., 1998). For example, these DNA diagnostic tests enable physicians to reassure patients with an initial incorrect clinical diagnosis of von Hippel–Lindau disease, to provide molecular, and thus more accurate diagnosis of von Hippel–Lindau disease, and to guide clinical management of members of families affected with *VHL*. Tests for somatic *VHL* mutations may be useful in classifying sporadic renal cancers. Although we encountered one refractory mutation using EndoV/Ligase scanning, this could be detected in a parallel PCR/LDR test.

The EndoV/Ligase mutation scanning method may become clinically useful for high throughput detection of somatic *p53* mutation in tumors. To date, various technologies used for pre-screening or screening are not sensitive enough, nor do they provide information about the precise *p53* mutation (Soussi and Béroud, 2001). Consequently, the majority of studies of mutational status of *p53* were performed on exons 5 through 8 in the common belief that only a few per cent of mutations were missed. In a recent study, we have demonstrated that more than 15% of mutations are found outside exons 5–8 (Soussi and Béroud, 2001). The most frequent missing mutations are located in exons 4, 10, and 9, respectively. Furthermore, the pattern of these mutations is different from those found in exon 5–8, suggesting that they could have an important impact on prognosis (Soussi and Béroud, 2001). Screening exon 4 to 10 with the EndoV/Ligase mutation scanning assay should lead to an improvement in the sensitivity of *p53* mutation detection. This may be further enhanced by performing the EndoV/Ligase mutation scanning assay on *p53* cDNA.

SNPs may serve as markers of disease, or may associate directly with the disease gene (Risch and Merikangas, 1996). Identification of SNPs and mutations in large stretches of DNA and in a large number of samples requires novel methods to replace classical ones, which were not designed for high-throughput screening. In a recent survey to discover new SNPs in 106 genes with 114 independent alleles, VDA was used alone or in combination with DHPLC. The approach was shown to be systematic and comprehensive with a discordancy rate of 21% compared to direct DNA sequencing (Cargill *et al*., 1999; Halushka *et al*., 1999). A larger scale survey of all exons and flanking intronic regions of 313 human genes identified 3899 SNPs in 82 samples using automated DNA sequencing (Stephens *et al*., 2001). Although these approaches are ideal when surveying a broad spectrum of the population for common variants or cosmopolitan SNPs (Stephens *et al*., 2001), identifying low frequency disease causing mutations with high penetrance will require screening dozens of candidate genes in hundreds of affected individuals and matched controls. Endo V/ligase mutation scanning supports fragment pooling 5–10-fold, as well as the ability to distinguish mutations in large fragments that would normally require four or more sequencing runs. This translates into a potential 10–40-fold increase in throughput. Confirming the precise nature of newly identified mutations can be achieved using automated sequencing, while the subsequent large-scale clinical studies to validate disease association can be facilitated by PCR/LDR/Universal Array (Gerry *et al*., 1999; Favis *et al*., 2000).

Materials and methods

Materials

All routine chemical reagents were purchased from Sigma Chemicals (St. Louis, MO, USA) or Fisher Scientific (Fair Lawn, NJ, USA). DNA sequencing kits, GeneScan-500 (TAMRA) Size Standard, and PCR kits were purchased from Applied Biosystems Division of Perkin-Elmer Corporation (Foster City, CA, USA). Deoxyribonucleoside triphosphate (dNTPs), bovine serum albumin (BSA), ATP, 7-deaza-dGTP were purchased from Boehringer-Mannheim (Indianapolis, IN, USA). TaqPlus Precision PCR kit was purchased from Stratagene (La Jolla, CA, USA). Proteinase K was purchased from QIAGEN (Valencia, CA, USA). Protein assay kit was from Bio-Rad (Hercules, CA, USA). Deoxyoligonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA, USA). Microcon 30 filters were purchased from Millipore (Bedford, MA, USA). DNA polymerase I (Klenow fragment) was purchased from New England Biolab (Beverly, MA, USA). Sep-Pak Cartridge C-18 was purchased from Waters (Milford, MA, USA). Centri-Sep^{TM} spin columns were purchased from Princeton Separation (Adelphia, NJ, USA). *Thermotoga maritima* Endonuclease V and *Thermus* species AK16D DNA ligase were purified as described (Huang *et al*., 2001; Tong *et al*., 1999).

For detecting K-*ras* mutations, genomic DNA was extracted from cell lines as described (Khanna *et al*., 1999). Cell lines HT29 and SW1417 contain the wild-type K-*ras* gene, while SW620 and SW480 contain the G12V (G→T) mutation. The ratio of wild-type to mutant sequence in the genomic DNA extracted from cell lines was determined using PCR/LDR: LS180,1 : 1.8 (wt: G12D, G→A); SW1116, 1 : 0.7 (wt: G12A, G→C) and HCT15, 1 : 1.1 (wt: G13D, G→A). Genomic DNA containing germline mutations, or tumor DNA containing sporadic mutations was extracted as described (Khanna *et al*., 1999). A blinded set of 50 adenocarcinomas comprised of a mix of stages I, II, III and IV were evaluated for p53 mutations using both PCR/LDR Universal Array (Favis *et al*., unpublished results) and EndoV/Ligase mutation scanning (this work). Tumors were frozen in liquid nitrogen directly after surgical resection. The tissue was not microdissected, i.e. five core samples were punched from various positions in the tumor and subjected to DNA extraction as described above.

PCR amplification

DNA sequences of PCR primers used in this study are listed in Table 2. 5′ end ((3′,6′-dippivaloylfluoresceinyl)-6-carboxamidohexyl)-1-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite (6-FAM) or 4,7,2′,7′-tetrachloro-(3′,6′-dippivaloylfluoresceinyl) - 6-carboxamidohexyl-1-O- (2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite (TET) labeled PCR primers were purified by separation on a 10% denaturing polyacrylamide gel (7M urea) as described (Applied Biosystems Inc. FC, CA, USA in 1992). PCR reactions (50 μl) contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 200 μM of each dNTP, 0.2 μM of each primer pair, 2.5 mM $MgCl_2$, 2.5 unit AmpliTaq DNA polymerase or AmpliTaq Gold DNA polymerase, and 100 ng genomic DNA. In the case of *VHL* exon 1, which has a high GC content, 2% dimethyl sulfoxide (DMSO) was included in the PCR reaction mixture. Addition of DMSO alone has been shown to improve amplification of some GC-rich sequences by disrupting base-paring (Pomp and Medrano, 1991). PCR amplification conditions for each fragment were as follows (gene, exon, polymerase, denaturation, cycles, cycling conditions): (K-*ras*, exon 1, AmpliTaq DNA polymerase, 94°C for 2 min, 30 cycles, 94°C for 15 s, 60°C for 2 min), (*BRCA 1* exon 2 and exon 20, *BRCA 2* exon 11 and *p53* exon 7, AmpliTaq Gold DNA polymerase, 95°C for 10 min, 35 cycles, 94°C 30 s, 60°C for 30 s, 72°C for 1 min), (*APC* exon 15, AmpliTaq Gold DNA polymerase, 95°C for 10 min, 30 cycles, 94°C for 30 s, 63°C for 75 s), (*VHL* exon 1 and 3, AmpliTaq DNA polymerase, 95°C for 2 min, 35 cycles, 94°C for 20 s, 66°C for 30 s, 72°C for 1 min), (*VHL* exon 2, AmpliTaq DNA polymerase, 95°C for 2 min, 35 cycles, 94°C for 20 s, 60°C for 30 s, 72°C for 1 min), (*p53* exons 5, 6 and 8, AmpliTaq DNA polymerase, 95°C for 2 min, 35 cycles, 94°C for 20 s, 65°C for 2 min).

For amplification of the 1.7 kb segment in *p53*, the PCR mixtures (50 μl) contained 1× TaqPlus Precision buffer, 200 μM dNTP, 0.2 μM of primers, 100 ng DNA template and 2.5 unit TaqPlus precision DNA polymerase (Stratagene, La Jolla, CA, USA). The PCR procedure included a predenaturation step at 95°C for 2 min, 35 cycles of two-step amplification with each cycle consisting of denaturation at 94°C for 20 s and annealing-extension at 68°C for 3 min. All PCR reactions above were completed with a final extension step at 72°C for 7 min.

Preparation of heteroduplexed DNA substrates

To remove *Taq* DNA polymerase, 4 μl of 20 mg/ml proteinase K (QIAGEN) was added to the PCR mixtures (50 μl) and incubated at 70°C for 10 min. Proteinase K was

inactivated by incubating at 80°C for 10 min. Amplicons containing wild-type sequence were added in approximately equal ratios when missing from the sample (i.e. pure mutant cell line DNA). The mixed PCR fragments, were heated at 94°C for 1 min to denature the DNA, and then cooled at 65°C for 15 min and at room temperature for 15 min to allow efficient formation of heteroduplex DNA.

EndoV/ligase mutation scanning assay

The reaction mixtures (20 μl) contained 10 mM N-(2-Hydroxyethyl)piperazine-N′-(2-ethanesulfonic acid) (HEPES) (pH 7.4), 1 mM dithiothreitol (DTT), 2% glycerol, 5 mM $MgCl_2$, 5% DMSO, 1–1.5 M N,N,N-trimethylglycine (betaine), 100 ng heteroduplexed PCR products, and 500 nM purified *Tma* Endonuclease V. The reaction mixtures were incubated at 65°C for 1 h. Fifteen μl of reaction mixtures from EndoV cleavage were supplemented with 2 μl of 10 × supplemental buffer (200 mM Tris-HCl, pH 8.5, 12.5 mM $MgCl_2$, 500 mM KCl, 100 mM DTT and 200 μg/ml BSA), 1 μl of 20 mM NAD^+, and 2 μl of 60 nM *Tsp*. AK16D DNA ligase. The resulting mixtures were incubated at 65°C for 20 min and terminated by adding equal volumes of GeneScan stop solution (50 mM ethylenediaminetertraacetic acid (EDTA), 1% blue dextran and 80% formamide). After denaturing at 94°C for 1 min, 3 μl of the mixtures were loaded onto a 6% acrylamide/Bisacrylamide (19 : 1) 0.2 mm thick denaturing gel containing 6M urea and electrophoresed for 1 h at 1000 V in TBE buffer at 45°C in an ABI-377 sequencer. Results were analysed using the Genescan program (Perkin Elmer).

Detection of mutations in long PCR fragment

For long range mutation scanning, the PCR fragments were filtered twice with Microcon 30 filters (10 mM Tris pH 7.5), prior to the Endo V cleavage reactions. This removed excess dNTP and primers while concentrating the sample twofold. After Endo V reactions, the fragments were washed in the same fashion to remove betaine and DMSO. The ligation reactions were carried out at 65°C for 20 min in the presence of 6 nM *Tsp*.AK16D ligase in the buffer containing 20 mM Tris-HCl (pH 8.5), 5 mM $MgCl_2$, 50 mM KCl, 10 mM DTT, 1 mM NAD^+ and 20 μg/ml BSA.

Acknowledgments

We thank Berton Zbar, Alan Friedman, Donald Bergstrom, Arnold Levine, David Thaler, Richard Cunningham, Daniel Notterman, Ken Offit and Alvaro Monteiro for helpful discussion and Norman Gerry, Carrie Shawber, Jing Lu, Jie Tong and Andrew Grace for their suggestions and technical assistance. We thank Yoke Kow for providing *E.coli* Endo V and helpful suggestions. We thank Laura Schmidt of the National Cancer Institute-Frederick Cancer Center for providing germline samples containing *VHL* mutations. We thank Harry Ostrer and Carol Oddoux from the Human Genetics Program at New York University Medical College for providing blinded genomic DNA samples from Ashkenazi Jewish individuals previously typed for the *APC*I1307K allele. We thank Steven Narod of the University of Toronto for providing samples containing the Ashkenazi Jewish *BRCA1* and *BRCA2* founder mutations. This work was supported by grants from the National Cancer Institute (P01-CA65930 and RO1-CA81467).

References

Abeliovich D, Kaduri L, Lerer I, Weinberg N, Amir G, Sagi M, Zlotogora J, Heching N and Peretz T. (1997). *Am. J. Hum. Genet.,* **60,** 505–514.

Ahrendt S, Halachmi S, Chow J, Wu L, Halachmi N, Yang S, Wehage S, Jen J and Sidransky D. (1999). *Proc. Natl. Acad. Sci. USA,* **96,** 7382–7387.

Aurelio ON, Cajot JF, Hua ML, Khwaja Z and Stanbridge EJ. (1998). *Cancer Res.,* **58,** 2190–2195.

Aurelio ON, Kong XT, Gupta S and Stanbridge EJ. (2000). *Mol. Cell. Biol.,* **20,** 770–778.

Beller U, Halle D, Catane R, Kaufman B, Hornreich G and Levy-Lahad E. (1997). *Gynecol. Oncol.,* **67,** 123–126.

Berman DB, Costalas J, Schultz DC, Grana G, Daly M and Godwin AK. (1996). *Cancer Res.,* **56,** 3409–3414.

Broll R, Stark A, Windhovel U, Best R, Strik MW, Schimmelpenning H, Schwandner O, Kujath P, Bruch HP and Duchrow M. (1999). *Eur. J. Cancer,* **35,** 1083–1088.

Bunz F, Hwang PM, Torrance C, Waldman T, Zhang Y, Dillehay L, Williams J, Lengauer C, Kinzler KW and Vogelstein B. (1999). *J. Clin. Invest.,* **104,** 263–269.

Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Lane CR, Lim EP, Kalayanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ and Lander ES. (1999). *Nat. Genet.,* **22,** 231–238.

Chen F, Slife L, Kishida T, Mulvihill J, Tisherman SE and Zbar B. (1996). *J. Med. Genet.,* **33,** 716–771.

Chen J and Thilly WG. (1994). *Environ. Health Perspect.,* **3,** 227–229.

Cotton RGH, Rodrigues NR and Campbell RD. (1988). *Proc. Natl. Acad. Sci. USA,* **85,** 4397–4401.

Dameron KM, Volpert OV, Tainsky MA and Bouck N. (1994). *Science,* **265,** 1582–1584.

Fahy E, Nazarbaghi R, Zomorrodi M, Herrnstadt C, Parker WD, Davis RE and Ghosh SS. (1997). *Nucleic Acids Res.,* **25,** 3102–3109.

Favis R, Day JP, Gerry NP, Phelan C, Narod S and Barany F. (2000). *Nat. Biotechnol.,* **18,** 561–564.

Fodde R and Losekoot M. (1994). *Hum. Mutat.,* **3,** 83–94.

Foster BA, Coffey HA, Morin MJ and Rastinejad F. (1999). *Science,* **286,** 2507–2510.

Gerry NP, Witowski NE, Day J, Hammer RP, Barany G and Barany F. (1999). *J. Mol. Biol.,* **292,** 251–262.

Giunta C, Youil R, Venter D, Chow CW, Somers G, Lafferty A, Kemper B and Cotton RG. (1996). *Diagn. Mol. Pathol.,* **5,** 265–270.

Guldberg P and Guttler F. (1994). *Nucleic Acids Res.,* **22,** 880–881.

Hacia J. (1999). *Nat. Genet. (Supplement),* **21,** 42–47.

Hacia JG, Brody LC, Chee MS, Fodor SP and Collins FS. (1996). *Nat. Genet.,* **14,** 441–447.

Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R and Chakravarti A. (1999). *Nat. Genet.,* **22,** 239–247.

Hansen LL, Justesen J and Kruse TA. (1995). *PCR Primer: a Laboratory Manual.* Diefenbach CW and Dveksler GS. Cold Spring Harbor Laboratory Press: New York pp. 275–286.

Haris II, Green PM, Bentley DR and Giannelli F. (1994). *Pcr Methods Appl.,* **3,** 268–271.

Hayashi K. (1991). *PCR Methods Appl.,* **1,** 34–38.

Heide I, Thiede C, Sonntag T, de Kant E, Neubauer A, Jonas S, Peter FJ, Neuhaus P, Herrmann R, Huhn D and Rochlitz CF. (1997). *Eur. J. Cancer,* **33,** 1314–1322.
Henke W, Herdel K, Jung K, Schnorr D and Loening SA. (1997). *Nucl. Acids Res.,* **25,** 3957–3958.
Huang J, Lu J, Barany F and Cao W. (2001). *Biochemistry,* **40,** 8738–8748.
Innis MA, Myambo KB, Gelfand DH and Brow MAD. (1988). *Proc. Natl. Acad. Sci. USA,* **85,** 9436–9440.
Khanna M, Park P, Zirvi M, Cao W, Picon A, Day J, Paty P and Barany F. (1999). *Oncogene,* **18,** 27–38.
Khrapko K, Hanekamp JS, Thilly WG, Belenkii A, Foret F and Karger BL. (1994). *Nucl. Acids Res.,* **22,** 364–369.
Korn SH, Moerkerk PT and de Goeij AF. (1993). *J. Clin. Pathol.,* **46,** 621–623.
Laken SJ, Petersen GM, Gruber SB, Oddoux C, Ostrer H, Giardiello FM, Hamilton SR, Hampel H, Markowitz A, Klimstra D, Jhanwar S, Winawer S, Offit K, Luce MC, Kinzler KW and Vogelstein B. (1997). *Nat. Genet.,* **17,** 79–83.
Libutti SK, Choyke PL, Alexander HR, Glenn G, Bartlett DL, Zbar B, Lubensky I, McKee SA, Maher ER, Linehan WM and Walther MM. (2000). *Surgery,* **128,** 1022–1027; discussion 1027–1028.
Liu J, He B, Qing H and Kow YW. (2000). *Mutat. Res.,* **461,** 169–177.
Liu Q and Sommer SS. (1995). *Biotechniques,* **18,** 470–477.
Makino R, Yazyu H, Kishimoto Y, Sekiya T and Hayashi K. (1992). *PCR Methods Appl.,* **2,** 10–13.
Mohr VH, Vortmeyer AO, Zhuang Z, Libutti SK, Walther MM, Choyke PL, Zbar B, Linehan WM and Lubensky IA. (2000). *Am. J. Pathol.,* **157,** 1615–1621.
Myers RM, Larin Z and Maniatis T. (1985). *Science,* **230,** 1242–1246.
Oddoux C, Struewing JP, Clayton CM, Neuhausen S, Brody LC, Kaback M, Haas B, Norton L, Borgen P, Jhanwar S, Goldgar D, Ostrer H and Offit K. (1996). *Nat Genet,* **14,** 188–190.
Oleykowski CA, Bronson Mullins CR, Godwin AK and Yeung AT. (1998). *Nucleic Acids Res.,* **26,** 4597–4602.
Perucho M, Forrester K, Almoguera C, Kahn S, Lama C, Shibata D, Arnheim N and Grizzle WE. (1989). *Cancer Cells,* **7,** 137–141.
Pomp D and Medrano JF. (1991). *Biotechniques,* **10,** 58–59.
Prives C and Hall PA. (1999). *J. Pathol.,* **187,** 112–126.
Rees WA, Yager TD, Korte J and von Hippel PH. (1993). *Biochemistry,* **32,** 137–144.
Ridanpaa M, Burvall K, Zhang LH, Husgafvel-Pursiainen K and Onfelt A. (1995). *Mutat. Res.,* **334,** 357–364.
Ridanpaa M and Husgafvel-Pursiainen K. (1993). *Hum. Mol. Genet.,* **2,** 639–644.
Risch N and Merikangas K. (1996). *Science,* **273,** 1516–1517.
Roa BB, Boyd AA, Volcik K and Richards CS. (1996). *Nat. Genet.,* **14,** 185–187.
Sarkar G, Yoon HS and Sommer SS. (1992). *Genomics,* **13,** 441–443.
Sgambati MT, Stolle C, Choyke PL, Walther MM, Zbar B, Linehan WM and Glenn GM. (2000). *Am. J. Hum. Genet.,* **66,** 84–91.
Soussi T and Béroud C. (2001). *Cancer,* **1,** 233–240.
Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han JH, Duan J, Carr JL, Lee MS, Koshy B, Kumar AM, Zhang G, Newell WR, Windemuth A, Xu C, Kalbfleisch TS, Shaner SL, Arnold K, Schulz V, Drysdale CM, Nandabalan K, Judson RS, Ruano G and Vovis GF. (2001). *Science,* **293,** 489–493.
Stolle C, Glenn G, Zbar B, Humphrey JS, Choyke P, Walther M, Pack S, Hurley K, Andrey C, Klausner R and Linehan WM. (1998). *Hum. Mutat.,* **12,** 417–423.
Struewing JP, Abeliovich D, Peretz T, Avishai N, Kaback MM, Collins FS and Brody LC. (1995). *Nat. Genet.,* **11,** 198–200.
Struewing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, Timmerman MM, Brody LC and Tucker MA. (1997). *N. Engl. J. Med.,* **336,** 1401–1408.
Suzuki Y, Orita M, Shiraishi M, Hayashi K and Sekiya T. (1990). *Oncogene,* **5,** 1037–1043.
Tabor S and Richardson CC. (1995). *Proc. Natl. Acad. Sci. USA,* **92,** 6339–6343.
Taylor GR. (1999). *Electrophoresis,* **20,** 1125–1130.
Thiagalingam S, Laken S, Willson JK, Markowitz SD, Kinzler KW, Vogelstein B and Lengauer C. (2001). *Proc. Natl. Acad. Sci. USA,* **98,** 2698–2702.
Tong J, Cao W and Barany F. (1999). *Nucl. Acids Res.,* **27,** 788–794.
Tortola S, Marcuello E, Gonzalez I, Reyes G, Arribas R, Aiza G, Sancho FJ, Peinado MA and Capella G. (1999). *J. Clin. Oncol.,* **17,** 1375–1381.
Underhill PA, Jin L, Zemans R, Oefner PJ and Cavalli-Sforza LL. (1996). *Proc. Natl. Acad. Sci. USA,* **93,** 196–200.
Walther MM, Reiter R, Keiser HR, Choyke PL, Venzon D, Hurley K, Gnarra JR, Reynolds JC, Glenn GM, Zbar B and Linehan WM. (1999). *J. Urol.,* **162,** 659–664.
Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ and Lander ES. (1998a). *Science,* **280,** 1077–1082.
Wang Y, Blandino G, Oren M and Givol D. (1998b). *Oncogene,* **17,** 1923–1930.
Webley KM, Shorthouse AJ and Royds JA. (2000). *J. Pathol.,* **191,** 361–367.
Winter E, Yamamoto F, Almoguera C and Perucho M. (1985). *Proc. Natl. Acad. Sci. USA,* **82,** 7575–7579.
Xu JF, Yang QP, Chen JY, van Baalen MR and Hsu IC. (1996). *Carcinogenesis,* **17,** 321–326.
Yan H, Kinzler KW and Vogelstein B. (2000a). *Science,* **289,** 1890–1892.
Yan H, Papadopoulos N, Marra G, Perrera C, Jiricny J, Boland CR, Lynch HT, Chadwick RB, de la Chapelle A, Berg K, Eshleman JR, Yuan W, Markowitz S, Laken SJ, Lengauer C, Kinzler KW and Vogelstein B. (2000b). *Nature,* **403,** 723–724.
Yao M, Hatahet Z, Melamede RJ and Kow YW. (1994). *Ann. NY Acad. Sci.,* **726,** 315–316.
Yao M and Kow YW. (1994). *J. Biol. Chem.,* **269,** 31390–31396.
Youil R, Kemper BW and Cotton RG. (1995). *Proc. Natl. Acad. Sci. USA,* **92,** 87–91.
Zirvi M, Nakayama T, Newman G, McCaffrey T, Paty P and Barany F. (1999). *Nucl. Acids Res.,* **27,** 40.
Zou Z, Gao C, Nagaich AK, Connell T, Saito S, Moul JW, Seth P, Appella E and Srivastava S. (2000). *J. Biol. Chem.,* **275,** 6051–6054.