Inverse PCR-Based RFLP Scanning Identifies Low-Level Mutation Signatures in Colon Cells and Tumors

Wei-Hua Liu, Manjit Kaur, Gang Wang, Penny Zhu, Yuzhi Zhang, and G. Mike Makrigiorgos

Department of Radiation Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts

ABSTRACT

Detecting the presence and diversity of low-level mutations in human tumors undergoing genomic instability is desirable due to their potential prognostic value and their putative influence on the ability of tumors to resist drug treatment and/or metastasize. However, direct measurement of these genetic alterations in surgical samples has been elusive, because technical hurdles make mutation discovery impractical at low-mutation frequency levels (<10\(^{-5}\)). Here, we describe inverse PCR-based amplified restriction fragment length polymorphism (iFLP), a new technology that combines inverse PCR, RFLP, and denaturing high-performance liquid chromatography to allow scanning of the genome at several thousand positions per experiment for low-level point mutations. Using iFLP, widespread, low-level mutations at mutation frequency 10\(^{-7}\)–10\(^{-5}\) were discovered in genes located on different chromosomes, e.g., OGG1, MSH2, PTEN, β-catenin, Bcl-2, P21, ATK3, and Braf, in human colon cancer cells that harbor mismatch repair deficiency whereas mismatch repair-proficient cells were mutation free. Application of iFLP to the screening of sporadic colon cancer surgical specimens demonstrated widespread low-level mutations in seven out of 10 samples, but not in their normal tissue counterparts, and predicted the presence of millions of diverse, low-incidence mutations in tumors. Unique low-level mutational signatures were identified for each colon cancer cell line and tumor specimen. iFLP allows the high-throughput discovery and tracing of mutational signatures in human cells, precancerous lesions, and primary or metastatic tumors and the assessment of the number and heterogeneity of low-level mutations in surgical samples.

INTRODUCTION

Large numbers of mutations are postulated to occur as early events in carcinogenesis (1, 2) and, for certain types of human tumors, to be a driving force in generating clonogenic, causative genetic changes required for multistage carcinogenesis (2–4). Cancers of the mutator phenotype are estimated to contain genetic alterations on the order of thousands per cell (2, 5). Widespread mutations occurring almost randomly at lower levels [mutation frequency (MF) <10\(^{-5}\)] throughout the genome are in fact estimated to have an even higher incidence (6) and possibly occur in a variety of human tumors (7). These putative low-level genomic events are assumed to take place very early in malignant transformation, and their presence is highly significant due to their potential use as molecular markers for early identification of genomic instability that can lead to cancer and their putative influence on the ability of tumors to resist drug treatment and/or metastasize (8, 9).

However, because such genetic events fall below the detection threshold of practically all mutation scanning methods, they have been very difficult to detect (7, 10). Phenotypic mutation detection methods applied to the identification of low-level nucleotide changes in the hypoxanthine phosphoribosyltransferase (HPRT) and lacI genes can, in fact, detect a very high number of low-level nucleotide changes in mismatch repair-deficient colon cancer human cell lines (11, 12) and in transgenic rat tumors (13). However, the inability of applying these mutation detection systems to nonclonogenic cells and to genes relevant to cancer has precluded direct examination of low-level mutations in surgical tumor samples that potentially contain repair defects similar to those observed in the cell lines. Due in part to this technical hurdle, the influence of low-level mutations in tumor onset, progression, and response to treatment has been difficult to assess (1, 14), and the origins of genomic instability in human cancer remain a subject of controversy (15).

We describe inverse PCR-based, amplified restriction fragment length polymorphism (iFLP), a new technique designed to allow high-throughput, genome-wide scanning for low-level mutations in cell lines and human tumors. DNA sequences are circularized after a self-ligation reaction and then exposed to a TaqI restriction enzyme digestion (Fig. 1A). DNA circles that do not normally contain a natural restriction site are converted to double-stranded linear DNA fragments only if they have acquired nucleotide changes leading to formation of TaqI site(s). By ligating TaqI-specific adaptors, these DNA fragments are PCR amplified selectively over their wild-type counterparts. Digestion by TaqI anywhere along the sequence leads to formation of a single DNA size that allows highly accurate and sensitive detection by size-resolving methods. We developed a system that harnesses this simple principle on a genome-wide scale by using high-throughput inverse PCR in conjunction with automated denaturing high-resolution DNA fragment size analysis. This technology, widespread, low-level mutations associated with genomic instability at the nucleotide level [single-nucleotide instability (SNI)] were discovered in repair-deficient colon cancer cells and colon cancer surgical specimens, and their extent and diversity were assessed. To our knowledge, this is the first report describing low-level (MF <10\(^{-5}\)) mutational signatures and their diversity in human solid tumor surgical specimens.

MATERIALS AND METHODS

Cell Culture, Surgical Samples, and DNA Extraction. The colorectal carcinoma cell lines SW-480 (ATCC CCL-226), DLD-1 (ATCC CCL-221) and HCT-116 (ATCC CCL-247) were purchased from the American Type Culture Collection. DLD-1 and HCT-116 cells are deficient in mismatch repair activity and display microsatellite instability (MSI) and a high spontaneous mutation rate at coding sequences (11, 16, 17). SW-480 cells do not display mismatch repair deficiency (18) or MSI (16, 19–21), but they harbor a G→A mutation in codon 273 of p53 gene (Arg→His substitution; Ref. 22). We observed that an NlsIII restriction site is generated in the mutated p53 sequence, and this property was used as a positive control in certain experiments. The osteosarcoma (HOS) cell line (ATCC CRL-1543), which is negative for mutations in both p53 and HPRT genes, was also used. After initial treatments, cell lines were grown for 32 generations in DME medium supplemented

Received 11/21/03; revised 1/8/04; accepted 1/21/04.

Grant support: NIH Grants 5 T32 CA09078 (W-H. Liu), CA083234 and The Joint Center for Radiation Therapy Foundation.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: W-H. Liu and M. Kaur contributed equally to this work.

Requests for reprints: G. Mike Makrigiorgos, Dana-Farber—Brigham and Women’s Cancer Center, Brigham and Women’s Hospital, Level L2, Radiation Therapy, 75 Francis Street, Boston, MA 02115. Phone: (617) 525-7122; Fax: (617) 632-6900; E-mail: mmakrigiorgos@partners.org.
with 10% fetal bovine serum (Life Technologies). To perform the HPRT assay, DLD-1 cells grown for 32 generations were treated with 6-thioguanine (6TG). 6TG, 6 μg/ml, was added along with DMEM and 10% fetal bovine serum after 2 h of cell settlement. Cells were allowed to grow for another ~12 doublings in the presence of 6TG to allow for phenotypic expression of the HPRT mutations or lack thereof. After this, cells were washed with PBS (Life Technologies), trypsinized with 0.25% trypsin-EDTA solution (Sigma), and counted before genomic DNA extraction. Ten human colon cancer surgical biopsies and corresponding normal tissue from the same patients were anonymously obtained from the Massachusetts General Hospital Tumor Bank. These specimens were removed immediately after surgery and snap-frozen in liquid nitrogen to preserve DNA. Genomic DNA was extracted from frozen samples using a commercial kit (QIAamp DNA mini kit; Qiagen, Inc.).

Single-Tube Procedure for Genome-Wide iFLP. A single-tube procedure was developed for genomic digestion via Msel, genome-wide circularization, exonuclease digestion, and TaqI-specific linker-ligation. Genomic DNA, 1–1000 ng depending on experiment, was suspended in 50 mM Tris-HCL, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, and 25 μg of BSA (pH 7.5) and digested with Msel (New England Biolabs, Beverly, MA) for 1.5 h to produce sticky ends for subsequent circularization. After heat inactivation of Msel (70°C for 30 min), DNA was circularized by addition of 1,200 units T4 DNA ligase (New England Biolabs) and incubation at 14°C overnight. After ligase inactivation (70°C for 30 min), circularized DNA was treated with 0.6 μl (12 units) of Escherichia coli exonuclease I (New England Biolabs) plus 2.4 μl (12 units) of A exonuclease (New England Biolabs) at 37°C for 1.5 h to eliminate noncircularized DNA. The temperature was then raised to 80°C to inactivate exonucleases. Circulized DNA was then digested with TaqI (New England Biolabs) followed by heat inactivation of the enzyme at 80°C for 40 min. In certain experiments, TaqI was substituted by the enzymes NlaIII, or HpyCH4IV. Linkers were then added to digested DNA, annealed at 50°C, and then slowly cooled to 10°C. Linkers used were: 5’-AGG CAA CTG TGC TAT CCG AGG GAA-3’ (2 μl from 2 μg/μl) and 5’-CGT TCC CTC GGA-3’ (2 μl from 1 μg/μl) for TaqI and HpyCH4IV; and 5’-AGG CAA CTG TGC TAT CCG AGC ATG-3’ (2 μl from 2 μg/μl) for NlaIII. This mixture was added 1 μl of T4 DNA ligase (2000 units/μl from New England Biolabs), followed by incubation at room temperature for 30 min. This single-tube procedure takes about 30 h to perform, including the overnight circularization.

In certain experiments, iFLP was performed on a genomic fraction rather than the full genome to enrich target DNA sequences and to allow higher selectivity in mutation detection. For this purpose, after the initial Msel digestion, DNA was separated on a 1% agarose gel and DNA in the molecular weight size region 400 to 700 bp was gel-purified (QIAquick gel extraction kit; Qiagen, Inc.). The iFLP procedure was subsequently performed with the purified enriched genomic DNA, which is estimated to contain about 5–10% of the original genome.

PCR Amplifications. After ligation in the single-tube protocol, 4–8 μl of the ligated genome were PCR-amplified using linker-specific primers 5’-AGG CAA CTG TGC TAT CCG AGG GAA-3’ for TaqI or HpyCH4IV-digested DNA and 5’-AGG CAA CTG TGC TAT CCG AGC ATG-3’ for NlaIII-digested DNA. PCR was carried out in a Perkin-Elmer GeneAmp PCR 9600 system (PE Biosystems) using a Titanium TaqPCR kit (Clontech) with thermocycling conditions: 72°C for 8 min followed by 22 cycles of 95°C for 30 s; 70°C for 1 min and 68°C for 5 min; and 4°C hold. A set of second, gene-specific quantitative PCR reactions (high-throughput inverse PCR) was then performed to detect presence of amplified inverted sequences, by using primers amplifying a short segment (~50–200 bp) from the genomic fragment of interest. For example, gene-specific PCR using the inverse primers 5’-AATGCGTGTTCCTCTTGACTG-3’ (forward) and 5’-TTCCGTCCCCAGTAGATTACC-3’ (reverse) results in a ~200-bp sequence. Presence of an amplicon following PCR indicates mutations leading to new restriction sites in a 611-bp circularized segment of p53, nucleotide 14361–nucleotide 14972, which was originally negative for restriction sites. To perform the high-throughput PCR reactions, a 0.5-μl sample from the first PCR product was used in each 10-μl reaction, using Titanium TaqDNA polymerase and the thermocycling conditions: 95°C for 1 min followed by 30 cycles of 95°C for 30 s; 68°C for 1 min and 68°C for 5 min; and 4°C hold.

Genomic regions selected for amplification via high-throughput PCR do not contain TaqI regions and therefore are not expected to yield a product unless a mutation modified the sequence to generate a TaqI digestion site. A 500-bp DNA fragment contains, on average, 80–100 mutable sites that can form a TaqI site following a single-nucleotide substitution, deletion, or insertion. Therefore each PCR reaction from the high-throughput PCR set evaluates 80–100 potential sequence changes simultaneously. In control experiments, use of an alternative enzyme whose recognition sequence is present in the wild-type sequence was used to serve as positive (+) control for iFLP, i.e., 100% mutant fragments. For example, a circularized DNA fragment with a HpyCH4IV restriction site in the wild-type sequence was digested with HpyCH4IV instead of TaqI yields a same-size PCR product with the one

---

**Fig. 1. Outline of the iFLP method. A, principle. B, high-throughput genome-wide iFLP.**
obtained following a TaqI-forming mutation. Conversely, some sequences that were examined were chosen to obtain no wild-type HpyCH41V while they also contained a wild-type TaqI site. These samples used the TaqI enzyme as a positive control and HpyCH41V as a mutation-scanning enzyme.

**Analysis of PCR Products and Estimation of MFs.** Inverse PCR products were analyzed either by ethidium bromide-stained 1% agarose gel electrophoresis or by HPLC chromatography on a WAVE dHPLC system (Transgenomics, Inc., Omaha, NE). The WAVE system is equipped with UV and fluorescent detectors and with two 96-well autosamplers that enable high-throughput analysis of PCR products. Eight μl of product from each PCR reaction were injected and analyzed at nondenaturing conditions, at 48°C. DNA elution profiles were recorded and displayed as chromatograms. To estimate the approximate MF for mutations discovered by iFLP, the positive iFLP controls were diluted into negative samples to produce samples with known ratios of mutant to wild-type alleles. High-throughput, inverse PCR was carried out using samples ofknown MF in parallel with unknown samples, and inverse PCR amplicons were quantitated via peak integration using the Navigator software of the WAVE system. The use of the high-sensitivity fluorescent marker allowed quantitation of PCR amplicons in the exponential phase of the PCR reaction, and therefore quantitation of MFs was not affected by PCR saturation effects. Experiments were repeated three independent times, and only PCR products demonstrating concordance among all three experiments were considered to be regions positive for TaqI-forming mutations.

**Verification and Sequencing.** Independent verification that iFLP-generated DNA amplicons correspond to genuine mutations at the corresponding centralized region was obtained by PCR amplifying directly from genomic DNA the indicated DNA region, followed by amplification via primer ligation at the mutation (23, 24). This approach uses the formation of a TaqI site at the mutation to ligate a linker and selectively PCR-amplify mutation-containing alleles. Amplification via primer ligation has been recently used to detect mutations with a selectivity of -10-3 (27) mutation-to-wild-type alleles (24). For verification of the iFLP-detected mutation at the *OGG1* gene, a 652-bp genomic DNA fragment encompassing the segment in which the mutation was detected by iFLP was amplified directly from genomic DNA, using primers 5'-AGGGGCAAATTAGGCTTACG-3' (forward) and 5'-AGGGCCAGATATTGTTACG-3' (reverse). One μg of genomic DNA was used in each PCR reaction using Advantage high-fidelity PCR amplification and touchdown PCR (94°C for 30 s; 10 cycles of 94°C for 20 s, 65°C for 20 s, and 68°C for 1 min, with the annealing temperature decreasing by 1°C/cycle; 20 cycles of 94°C for 20 s, 55°C for 20 s, and 68°C for 1 min; 6°C for 6 min; 4°C hold). Gel-purified PCR products (50 μg) were digested with TaqI enzyme, purified by QIAquick PCR purification kit (Qiagen, Inc.), and ligated with linkers 5'-AAGCTGCTATATTGGGATCCGGC-3' (forward) and 5'-GCTAGCTATATTGGGATCCGGC-3' (reverse). Each PCR product was gel-purified to remove traces of genomic DNA (QiAquick gel extraction kit; Qiagen, Inc.), and DNA was quantitated with picogreen (Molecular Probes). The iFLP single-tube protocol was then applied to the amplified DNA fragment. To verify that single-gene iFLP can detect the single *NalIII* restriction site present in SW-480 cells, which is otherwise absent in the wild-type sequence, *NalIII* digestion was used instead of TaqI.

**Detection of Mutations in HPRT.** Using single-gene iFLP, a 550-bp region from the *HPRT* gene was amplified and tested via iFLP for mutations forming *NalIII* restriction sites. The primers used were 5'-TTT AAC GTC AGT CCC TCT CTC T-3' (forward) and 5'-CCA GTT CTA AGG AGG TCT GTA-3' (reverse), and the PCR thermocycling conditions were the same as for the p53 region described above. The amplified HPRT fragment does not contain *NalIII* sites, however we observed that a low-level mutation generating an *NalIII* site (C→T missense mutation, C3983T, in codon 508 of the *HPRT* gene exon 7, in GenBank sequence no. M26434) was previously reported in DLD-1 mismatch repair-deficient cells (17, 21). To evaluate iFLP detection of this mutation and to quantitate the detected MF, the amplified HPRT fragment was mutagenized using mutagenesis via overlap extension (25), and a mutation-positive DNA fragment was obtained. The mutation-positive sample was diluted into negative samples to produce mixed samples with decreasing ratio of mutant-to-wild type alleles to a MF of 1 × 10^-6. Ethidium bromide-stained agarose gels of the PCR products were quantitated in a gel-detection system (α-Innotech Corp.) equipped with a low-level light-detecting camera and software for quantitative determination of fluorescent gel bands. The reproducibility of single-gene iFLP and genome-wide iFLP was verified by at least three independent experiments.

**MSL.** Experiments for determination of MSL were performed using the dHPLC method (26, 27) for assessment of BAT25 and BAT26 microsatellite markers (28). The procedures described by Pan et al. (26) were followed with minor modifications to amplify PCR fragments containing the microsatellite markers BAT25 and BAT26 and assess fragment length changes via the dHPLC. In brief, primers used for BAT26 were: 5'-GCAAGTCAGAGCCTTT-AAAC-3' (forward) and 5'-CCATACACACTTTTATCCC-3' (reverse); and primers for BAT25 were: 5'-GGGAGATGATTTCTAAAGAG-3' (forward) and 5'-CTTAAATCTGTCTCTTCTC-3' (reverse). A touchdown PCR reaction using titanium Taq polymerase (BD-Clontech) was used: 94°C for 30 s; 10 cycles of 94°C for 20 s, 65°C for 20 s, and 68°C for 1 min, with the annealing temperature decreasing by 1°C/cycle, touchdown PCR: 20 cycles of 94°C for 20 s, 55°C for 20 s, and 68°C for 1 min; 6°C for 6 min; 4°C hold. PCR products were injected in the dHPLC at 50°C using dHPLC gradient conditions described by Pan et al. (26), and presence of MSL for the specific marker was detected as a shift in the obtained dHPLC peak.

**RESULTS**

**Single-Gene iFLP.** Generation of an *NalIII* site by a single-point mutation anywhere along a circularized DNA sequence produces a PCR amplicon of a single size in the iFLP assay. This facilitates detection using size-separation visualization methods (gel electrophoresis, dHPLC) and distinguishes mutations from PCR mis-priming events. Single-gene iFLP was carried out in a p53 region that contains a G→A mutation generating a new *NalIII* site in SW-480 cells, as well as in HOS cells that are wild type and contain no *NalIII* site. A strong gel band appears at ~600 bp for cells carrying the mutation, which is absent for wild-type cells (Fig. 2A, Lanes 1 and 7, respectively). The amplified PCR product was excised from the gel and sequenced, and the expected mutation was confirmed. To examine the selectivity of mutation detection (i.e., the minimum mutant-to-wild type alleles detectable), gradual dilutions of the PCR product from mutation-positive genomic DNA into wild-type DNA were used (Fig. 2A, Lanes 2–6). The mutation was detectable within a million-fold excess wild-type DNA. SW-480 cells were then examined for mutations generating *NalIII* sites in a second DNA region, within the *HPRT* gene, which does not contain *NalIII* site in the wild type. No agarose-resolved bands were evident (Fig. 2B, Lane 6). In the presence of *NalIII* site, the anticipated iFLP amplicon is 558 bp. An HPRT DNA fragment mutagenized to contain *NalIII*-forming mutation
Lanes 1–6, MF of \(10^{-2}\), \(10^{-4}\), \(10^{-6}\), \(10^{-8}\), and 0, respectively. A, similar to A, but iFLP was applied for an HPRT gene fragment. Lanes 1–7, MF of \(10^{-2}\), \(10^{-3}\), \(10^{-4}\), \(10^{-5}\), \(10^{-6}\), and 0, respectively. B, similar to A, but iFLP was applied for an HPRT gene fragment. Lanes 1–3, SW-480 cells (mutation negative), 6-TG-selected DLD-1 cells, and mutation-positive control, respectively. D, single iFLP applied to the HPRT gene for non-6TG-treated DLD-1 or SW-480 cells, and dilutions of mutation positive into mutation negative. Lanes 1–7, SW-480, DLD-1, MF of \(10^{-6}\), \(10^{-3}\), \(10^{-4}\), \(10^{-5}\), and \(10^{-7}\), respectively.

(C→T missense mutation in codon 508) was then gradually diluted within the wild-type HPRT fragment and a \(~550\)-bp band with an intensity inversely proportional to the MF was observed, in the presence of up to a million-fold excess wild-type DNA. Given that there are 87 mutable positions that can generate NlaIII sites via a single-base substitution, insertion, or deletion on the examined HPRT fragment, these data indicate that single-gene iFLP scans the circularized HPRT fragment at 87 positions for specific base changes, at a selectivity of \(~10^{-6}\) at each position.

**Scanning for HPRT Mutations in Repair-Deficient Cells.** To examine whether single-gene iFLP can detect random, low-level mutations generated at the HPRT gene by genomic instability, the mismatch repair-deficient cell line DLD-1 was cultured for 32 generations so that development of previously reported base changes in HPRT can be anticipated, including an NlaIII-mutation forming at codon 508 (17, 21). Initially, a 6TG treatment was applied to select for HPRT mutants, and genomic DNA extracted from such mutants was examined via single-gene iFLP at HPRT. A strong 558-bp gel band was resolved when 6TG-treated DLD-1 mutants were screened (Fig. 2C, Lane 2). Excision of this band from the gel followed by sequencing revealed the anticipated C→T transition mutation that forms a new NlaIII site. Next, DLD-1 and SW-480 cells were examined in the absence of 6TG treatment. The 558-bp band was absent for the repair-proficient SW-480 cells but present for the repair-deficient DLD-1 cells, albeit with a ~5-fold lower intensity than when DLD-1 were 6TG-treated (Fig. 2D, Lanes 1 and 2, respectively), as determined by microdensitometry performed on the two gel pictures using a common standard. After excision of this gel band, the DNA was sequenced, and the anticipated mutation was confirmed. Semiquantitation of the MF revealed in DLD-1 cells was performed. By including dilutions of an artificially mutagenized HPRT fragment at codon 508 within the wild-type HPRT fragment and a /H11011 3 T missense mutation in codon 508) was then gradually diluted with an intensity inversely proportional to the MF was observed, in the presence of up to a million-fold excess wild-type DNA. Given that there are 87 mutable positions that can generate NlaIII sites via a single-base substitution, insertion, or deletion on the examined HPRT fragment, these data indicate that single-gene iFLP scans the circularized HPRT fragment at 87 positions for specific base changes, at a selectivity of \(~10^{-6}\) at each position.

**Genome-Wide iFLP.** To apply iFLP to high-throughput mutation screening, we tested whether the single-tube procedure can detect the NlaIII-forming G→A missense mutation in the p53 gene of SW-480 cells. Genomic DNA extracted from HOS (mutation-negative) cells and SW-480 cells were subjected to the single-tube iFLP protocol to examine presence of mutation-generated NlaIII sites on circularized segments from p53. After inverse PCR, strong bands appeared on ethidium gels for the mutant SW-480 (Fig. 3A, Lanes 2, 4, 6, 8, and 10) but not for the wild-type HOS cells. Variation of the amount of starting genomic DNA (1–200 ng) demonstrated that iFLP applied on 1 ng of genomic DNA is sufficient to generate a strong signal (Fig. 3A, Lane 10). A particularly strong gel band was obtained when 200 ng of genomic DNA enriched for the examined sequences (400–700-bp fraction of Msel-digested genome) was used as the starting material in the single-tube procedure (Fig. 3A, Lane 4). Enrichment allows the highest selectivity in iFLP mutation detection, because it ensures the presence of several copies of mutated sequences in the sample even at low MFs, \(~10^{-5}\)-mutant-to-wild type ratio.

![Fig. 2. Application of single-gene iFLP in detecting genomic instability-produced mutations. A, selectivity of iFLP for a p53 fragment, dilutions of mutation-positive (SW-480) into negative (HOS) DNA. Lanes 1–7, MF of \(10^{-2}\), \(10^{-3}\), \(10^{-4}\), \(10^{-5}\), \(10^{-6}\), \(10^{-7}\), and 0, respectively. B, similar to A, but iFLP was applied for an HPRT gene fragment. Lanes 1–6, MF of \(10^{-2}\), \(10^{-4}\), \(10^{-6}\), \(10^{-8}\), and 0, respectively. C, single-gene iFLP applied to the HPRT gene for 6TG-treated DLD-1 cells. Lanes 1–3, SW-480 cells (mutation negative), 6 TG-selected DLD-1 cells, and mutation-positive control, respectively. D, single iFLP applied to the HPRT gene for non-6TG-treated DLD-1 or SW-480 cells, and dilutions of mutation positive into mutation negative. Lanes 1–7, SW-480, DLD-1, MF of \(10^{-6}\), \(10^{-3}\), \(10^{-4}\), \(10^{-5}\), and \(10^{-7}\), respectively.](Image 2x)
Genome-Wide iFLP Screening of Mismatch Repair-Deficient Colon Cell Lines and Surgical Colon Specimens. Genome-wide iFLP was first applied to screening 30 DNA segments from a number of genes for widespread mutations in the mismatch repair-deficient cell lines DLD-1 and HCT-116 and the repair-proficient SW-480 and analyzed by gel electrophoresis. Inverse PCR products for DLD-1 and SW-480 cells along with mutation-positive controls (+) for each cell line are depicted in Fig. 3B. Out of 30 regions screened, four regions belonging to diverse genes demonstrated formation of new restriction sites for DLD-1 cells (Fig. 3B, arrows), whereas all 30 regions from SW-480 cells were negative.

Next, genome-wide iFLP was combined with automated dHPLC detection and applied to analyze a total of 48 gene regions using genomic DNA from the colon cancer cell lines DLD-1, HCT-116, and SW-480, as well as from 10 colon cancer and corresponding normal tissue surgical samples. The results confirmed the gel electrophoresis findings for the DLD-1 and SW-480 cell lines (Fig. 4). Thus, chromatograms from the OGG1, BRCA2, and MSH2 regions of interest from DLD-1 cells showed small peaks at the same retention time as the mutation-positive controls, whereas SW-480 cells showed no peaks. To estimate MFs for the identified mutations, dHPLC was used in the high-sensitivity fluorescence mode, and experiments using dilutions of the mutation-positive controls (obtained by digesting the circularized DNA with an enzyme containing a wild-type restriction site) into negative samples were performed together with the DLD-1 and SW-480 samples. Fig. 4E depicts a typical dilution experiment carried out for the BRCA2 region for HCT-116 and SW-480 cells, from which a MF $\sim 2 \times 10^{-4}$ was derived for HCT-116 cells. This approach also indicated that the detection limit of the method is approximately $10^5$ mutant: wild type ratio (Fig. 4E). Therefore, for starting genomic DNA material of 500 ng, which contains about $10^5$ genomic copies, the lowest MF detectable is MF $\sim 10^{-4}$ to $10^{-5}$. Similarly conducted dilution experiments indicated that, for the OGG1 gene in DLD-1 cells, the MF observed was MF $= 6 \times 10^{-3}$ (average of three independent experiments). Confirmation and identification of mutations in the OGG1 gene of DLD-1 cells was obtained by performing an independent mutation assay that PCR-amplifies the mutated sequence via ligation of a primer at the mutation (24), as described in “Materials and Methods.” Sequencing of the mutated fragment identified a Taq1-forming G$\rightarrow$A mutation in the coding region of OGG1 (codon 346; GenBank accession no. Y13277; missense Gly$\rightarrow$Arg).

Similar to results obtained with repair-deficient colon cancer cell lines, dHPLC screening of colon tumor, and corresponding normal tissue surgical samples resulted to PCR products present in the tumor and absent in the corresponding normal tissues. In Fig. 5, four representative results from tumor/normal samples 1–4 are depicted, together with dilutions of the positive controls that indicate the expected position of Taq1-forming mutations. By using serial dilutions of the mutation-positive controls, MFs for the 48 gene regions were subsequently derived. In Fig. 6, typical results are plotted for the mismatch repair-deficient cells DLD-1 and HCT-116, as well as for four colon tumor samples. The overall iFLP-identified mutation rates for DLD-1 regions (MF$\sim 2\sim 9 \times 10^{-3}$) are generally higher than for HCT-116 regions (MF$\sim 1\sim 8 \times 10^{-4}$), and the two cell lines contain mutations at distinct sequences. In addition, the total number of DLD-1 mutations in the 48 regions examined (sum of each mutation-positive segment $\times$ its own MF) is about 10-fold higher than for HCT-116. Bhattacharyya et al. (17) observed a similar relation among MFs in DLD-1 (MF$\sim 4 \times 10^{-3}$) and HCT-116 (MF $\sim 6 \times 10^{-4}$) and found distinct mutation spectra among the two cell lines by using the HPRT phenotypic assay. Therefore the iFLP data indicate that the genomic instability previously observed in HPRT is indicative of “global instability” occurring throughout the DLD-1 and HCT-116 genomes, as hypothesized (11, 12, 17), and that widespread, low-level mutations underlie the presence of a mutator phenotype in these cell lines. The difference in repair defects among the two cell lines may explain the differences in the mutation spectra. Mutation spectra depicted in Fig. 6 for the four colon tumor specimens contain regions that appear commonly mutated across most of the samples (e.g., MSH6, PTEN_C), indicating that these may be “hotspots” for the generation of mutations. In general however, the low-level mutational spectrum is unique to the particular tumor sample and cell line. Overall, by examination of 48 gene regions via dHPLC, an extensive pattern of Taq1-forming mutations was observed in seven out of 10 samples, whereas three samples contained no iFLP-detectable mutations (Table 1). These data indicate widespread genomic instability in sporadic colon tumors that manifests as numerous low-level mutations in gene-coding and noncoding regions. To estimate the overall number of diverse low-level mutations that would be discovered if iFLP was applied throughout the genome, the mean number of tumor-specific mutations per region examined (e.g., for tumor 1, 11 mutations in 48

![Diagram](Image68x374 to 272x748)
regions, i.e., ~0.23 mutations per ~500 bp) is scaled by the size of the human genome (3 × 10^9 bp), yielding ~1.4 × 10^6 distinct low-level mutations. The overall number of single-point mutations is probably larger, because only a fraction of all possible nucleotide changes generate TaqI sites and can be detected by iFLP.

**MSI.** The MSI status of the colon cancer cell lines and the 10 paired colon samples was examined by screening the mononucleotide markers BAT25 and BAT26. Changes in mononucleotide repeat size manifested as significant shifts in the obtained chromatograms (Fig. 7). Samples were classified as MSI-H if both BAT25 and BAT26 were shifted in the tumor relative to the normal colon samples, MSI-L if one of the two markers was shifted, and MSI-S if none of the two markers were shifted. Assessment of MSI status in this manner was recently shown (26) to be in excellent agreement to the results obtained by the Bethesda guidelines using multiple microsatellite markers. Repair-deficient cell lines DLD-1 and HCT116 demonstrated mononucleotide repeat changes for both BAT25 and BAT26 relative to the repair-proficient cell line SW480 (data not shown) and were thus classified as MSI-H. Two of 10 colon tumors, samples 1 and 5, demonstrated significant shifts in the dHPLC chromatograms compared with their normal tissue counterparts on both BAT26 and BAT25 and were classified as MSI-H. The remaining eight colon

![Fig. 5. Colon cancer and normal tissue surgical samples examined via genome-wide iFLP. Typical chromatograms are depicted from four gene segments indicating mutations in four colon cancer/normal tissue specimens (chromatograms A–D, gene segments PTEN, VEGF, CHK2, and β-catenin, respectively). Mutation-positive controls (+) for each gene segment diluted 100-fold for A, B, and D and 1000-fold for C and run under same conditions to indicate the retention times expected if mutations are present are also depicted. Experiments were repeated three times.](image)

![Fig. 6. iFLP-derived mutational signatures (MF versus gene segment) for repair-deficient colon cancer cell lines DLD-1 and HCT-116 and four colon surgical specimens for 30 of the 48 regions examined. (The remaining 18 regions were negative for the colon samples). The error bars represent SDs from three independent experiments. The list of gene segments carrying the mutations is depicted on the right. For genes in which more than one region was examined, these were indicated by A, B, or C.](image)

| Table 1 Percentage of DNA segments containing low-level mutations SNI for colon cancer cell lines and surgical samples, along with their MSI status |
|---------------------------------|-----------------|
|                                | SNI             | MSI status |
| DLD1 cells, repair deficient    | 13/48 (27%)     | MSI-H      |
| HCT116 cells, repair deficient  | 5/48 (10%)      | MSI-H      |
| SWA480 cells, repair proficient| 0/48 (0%)       | MSI-S      |
|Colon sample                    | Normal          | Tumor      |
|Sample 1                        | None            | 11/48 (23%)| MSI-H |
|Sample 2                        | None            | 16/48 (34%)| MSI-S |
|Sample 3                        | None            | 8/48 (17%) | MSI-S |
|Sample 4                        | None            | 6/48 (13%) | MSI-S |
|Sample 5                        | None            | 27/48 (56%)| MSI-H |
|Sample 6                        | None            | 0/48 (0%)  | MSI-S |
|Sample 7                        | 1 (2%)          | 4/48 (8%)  | MSI-S |
|Sample 8                        | None            | 0/48 (0%)  | MSI-S |
|Sample 9                        | None            | 0/48 (0%)  | MSI-S |
|Sample 10                       | None            | 14/48 (29%)| MSI-S |
|Overall                         | 7/10 (70%)      | 2/10 (20%) |
tumor samples were negative for both microsatellite markers and were classified as MSI-S. Table 1 summarizes the MSI results.

**DISCUSSION**

Carcinomas are generally heterogeneous and intrinsically redundant populations of cells generated by transformations including diverse mutations in individual genes and the loss or gain of entire chromosomes (aneuploidy; Ref. 8). To this end, diverse and widespread low-level mutations are expected to have profound implications in the ability of tumors to resist cytotoxic drug treatment (8). For example, specific mutations conferring resistance to treatment by the drug STI-571 in chronic myelogenous leukemia are known to preexist as low-level mutations in the untreated tumor (29–32). After selection by the drug treatment, some of these low-level mutations become causative, prevalent genetic changes (29, 33–35). Accordingly, assessment of the extent and diversity of low-level mutations in tumors would be very useful for estimation of the likelihood of drug resistance and tumor progression (8).

iFLP provides a methodology for achieving this purpose in surgical specimens, and demonstrates the existence of low-level mutations in sporadic colon cancers at MF \(10^{-2} \text{–} 10^{-4}\). DNA fingerprinting methods, arbitrarily primed (AP)-PCR, and inter-simple-sequence (SS) PCR, have also been applied to the identification of mutations in colon cancer surgical specimens (5, 36, 37). However, because these methods do not detect mutations at levels below MF \(10^{-1} \text{–} 10^{-2}\), low-level mutations are not detected, and the incidence of potential resistance-causing mutations is probably underestimated. This is illustrated in Fig. 8, in which the number of diverse mutations detected in Fig. 6 in the colon cancer samples is replotted versus their MF and indicates that mutations scored decrease rapidly as the selectivity of mutation detection decreases. For example, for tumor 1, a method with a selectivity of \(10^{-3}\) mutant-to-wild type alleles would detect seven mutations, whereas another method with a selectivity of \(10^{-2}\) would only detect one mutation. Therefore, to effectively estimate the heterogeneity and number of mutations in tumors, one must detect the diverse low-level mutation load of a surgical specimen, which is directly dependent on the selectivity of the detection method.

Our results show similar numbers of low-level mutations in colorectal cell lines and a substantial fraction of colon tumors and indicate that millions of diverse mutations are present in the majority of sporadic colon carcinomas examined. Although most of these mutations are probably not functional (2), their presence is indicative of an ongoing genomic instability at the single-nucleotide level (SNI), which may induce mutations with the potential to affect tumor development and progression. A similar SNI has been observed in cell lines and transgenic animals using phenotypic mutation assays (38, 39). MSI (21, 40) and chromosomal instability (41) are two established forms of genomic instability in colorectal cancer and other tumors (9, 10). However, many tumors display neither MSI nor chromosomal instability but different forms of genomic instability (42–44). For example, as demonstrated in Table 1, SNI of the form detected by iFLP often exists in the absence of MSI (38, 39), indicating that different repair defects may be responsible for each type of instability. Unlike MSI that only affects those protein-coding regions containing repeat sequences amenable to instability, SNI may potentially impact every coding sequence, thereby actively influencing processes that lead to tumor formation. Accordingly, it is possible to use iFLP to clarify types and subtypes of surgical specimens with MSI, the coexistence or absence of MSI and chromosomal instability and how early in the tumor formation process, such low-level point mutations build-up. It is also evident from Fig. 6 that, when SNI is significant, a single iFLP experiment can efficiently identify numerous low-level mutations and their frequencies, allowing generation of a mutational signature for each sample. These signatures can potentially be used for matching metastatic lesions with primary tumor sites or for detecting tumor recurrence by tracing the tumor signature in surgical margins or bodily fluids (9).

The described iFLP fills a current need in mutation detection methodology, i.e., the identification of unknown, random point mutations at low MFs. As Fig. 6 indicates, distinct sequences become mutated in DNA depending on the specific repair deficiency leading to instability (45, 46), thus it is difficult to predict a priori where a mutation is likely to occur to apply techniques for known mutation detection. Thus, despite development of RFLP-based methods that detect mutations at specific sequence positions down to a MF of \(10^{-7}\) (47, 48), random, low-level mutations generated by mismatch repair-induced genomic instability have proven hard to identify (49). Accordingly, mutation scanning rather than RFLP-based mutation detection at known sequence positions is required. On the other hand, PCR-based mutation scanning methods such as DGGE, dHPLC, and CDCE are ultimately limited by PCR errors (50, 51) and often have to be combined with phenotypic selection systems to

---

**Fig. 7.** dHPLC-based detection of MSI. Representative shifts in the chromatograms between tumor and normal tissue are depicted for the mononucleotide marker BAT26 in colon samples 1 and 3. a.u., arbitrary units.

**Fig. 8.** Influence of the selectivity of mutation detection on the numbers of distinct mutations identified in surgical samples. The data of Fig. 6 were replotted for the four colon tumor samples to illustrate the number of mutated DNA segments that would be identified if the mutation detection selectivity varied as designated on the X axis.
acquire the necessary selectivity (52). iFLP is not affected by polymerase errors and allows mutation detection at thousands of sequence positions per experiment combined with the high selectivity of RFLP at each position. The number of sequences examined can be scaled at will by adjusting the number of inverse PCR reactions performed in the last step. By magnifying the effective “target size,” iFLP multiplies the capabilities of RFLP detection by the total number of mutable sites, and a limited mutation scanning occurs that increases the probability of detecting widespread, low-level mutations, i.e., a hallmark of genomic instability (2, 53). It is estimated that about 3–5% of all possible mutations in each sequence can generate the TaqI recognition sequence via a single-nucleotide change and therefore can be identified by iFLP. By replacing TaqI with different restriction enzymes, more sequence changes can be detected. Because the assay can be applied from nanogram amounts of genomic DNA (Fig. 3A), it should be possible to perform high-throughput mutation scanning from minute biopsies obtained by needle aspiration or laser capture microdissection. By using methylation-sensitive enzymes, it should be possible to adapt iFLP to the discrimination of methylated from nonmethylated DNA sequences. Additional envisioned uses of iFLP include evaluating tissues for carcinogenic exposures that generate random, low-level mutations and identifying genetic variations in pooled DNA from animals or plants.

ACKNOWLEDGMENTS

We acknowledge the assistance of Mohamet Miri and Dr. Frank Haluska in obtaining tissue specimens from the Massachusetts General Hospital Tumor Bank and of the Cooperative Human Tissue Network.

REFERENCES

Inverse PCR-Based RFLP Scanning Identifies Low-Level Mutation Signatures in Colon Cells and Tumors

Wei-Hua Liu, Manjit Kaur, Gang Wang, et al.

Cancer Res 2004;64:2544-2551.

Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/64/7/2544

Cited articles
This article cites 52 articles, 26 of which you can access for free at:
http://cancerres.aacrjournals.org/content/64/7/2544.full#ref-list-1

Citing articles
This article has been cited by 2 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/64/7/2544.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.