Precise assessment of microsatellite instability using high resolution fluorescent microsatellite analysis

Shinya Oda$^{1, *}$, Eiji Oki$^{1}$, Yoshihiko Maehara$^{1}$ and Keizo Sugimachi$^{1,2}$

$^{1}$Cancer Center, Kyushu University Hospital and $^{2}$Department of Surgery II, Faculty of Medicine, Kyushu University, Fukuoka 812-82, Japan

Received June 3, 1997; Revised and Accepted July 21, 1997

ABSTRACT

The instability of microsatellite sequences dispersed in the genome has been linked to a deficiency in cellular mismatch repair. This phenotype has been frequently observed in various human neoplasms and is regarded as a major factor in tumorigenesis. To demonstrate alterations in microsatellite sequences, polymerase chain reaction (PCR) and electrophoretic analysis are techniques often used. However, the electrophoretic profiles of PCR-amplified microsatellite sequences have not been well characterized. Moreover, the conventional method using autoradiography has critical problems in detection characteristics and migration accuracy. We made use of fluorescence-labeled PCR and laser scanning with linear detection characteristics, so as to detect bands quantitatively. Next, we characterized Taq polymerase-dependent modification of the amplified microsatellite sequences, using artificially synthesized microsatellite alleles and we optimized the electrophoretic profiles by enzymatic modification with T4 DNA polymerase. We developed a dual fluorescence co-electrophoresis system, in which both samples derived from cancer and normal tissues are electrophoresed in the same lane, in order to minimize migration errors. These improvements remarkably facilitate precise and objective assessments of microsatellite instability. Analyzing many positive cases in cell lines and tissue specimens, we classified all the patterns of microsatellite alteration and set up new criteria for assessing microsatellite instability.

INTRODUCTION

DNA mismatch repair efficiently eliminates the mutagenic mismatch of nucleotides in the genome (1,2). As demonstrated in hereditary non-polyposis colon cancer (HNPPCC) (3,4), the defect of mismatch repair correlates with a high risk of carcinogenesis especially in the digestive tract (5–8). Sufficient data have accumulated to suggest that deficiency of mismatch repair also occurs in other human neoplasms. In this context, investigation of this abnormality is of great significance in clinical diagnostics, which ultimately will lead to assessment of the risk of newly or second occurring malignancies. Alteration of dinucleotide repeats in microsatellite sequences, termed microsatellite instability (MSI) or replication error (RER), is used as a diagnostic criterion of mismatch repair deficiency, since the mismatch repair system also repairs dinucleotide repeats looping out of the strand, events possibly generated by slippage of replicational polymerases. Methodological problems of conventional methods analyzing this alteration have remained.

In microsatellite analysis, sequences including dinucleotide repeats are usually amplified by polymerase chain reaction (PCR). The band patterns obtained by electrophoresis of the amplified products can be complicated, one reason for the complexity being that microsatellite sequences on paternal and maternal alleles are simultaneously amplified and the products often cross during electrophoresis. Second, the amplified products often consist of multiple fragments of different lengths. This phenomenon may not be derived only from heterogeneity of genomic templates, since Taq polymerase may possibly slip on a short stretch of dinucleotide repeat, as do the replicational polymerases. In addition, Taq polymerase has terminal deoxyribonucleotidyl transferase activity to add an extra one nucleotide at the 3′ end of PCR products. This activity of Taq polymerase presumably exerts great influence on the complexity of the amplified products. Moreover, polyacrylamide gel electrophoresis is susceptible to migration errors and autoradiography is prone to detection errors. Therefore, judgment of microsatellite instability based on presently used methods leaves much to be desired.

New techniques for nucleic acid electrophoresis, using fluorescence labeling and laser scanning, have been developed. In some systems, each nucleic acid fragment is quantitatively detected and mobility of the fragments is standardized with accuracy of one base pair. Although some investigators have already applied this fluorescence scanning for microsatellite analysis (9–11), problems in electrophoretic profiles of microsatellite sequences and migration accuracy have not been given due attention. We examined several polymerases harboring 3′–5′ exonuclease activity to reduce the complexity of the amplified microsatellite sequences and we determined the basic profile of the amplified microsatellite sequences optimized for microsatellite instability analysis. To minimize migration errors we examined...
the electrophoretic mobility and specific intensity of fragments labeled with various fluorescence compounds and determined the combination and the loading ratio in which the phases and the signal magnitudes of two electrophoretic profiles match completely. These improvements made feasible assessments of microsatellite instability while avoiding complicated electrophoretic profiles, migration inaccuracy or lack of a quantitative detection.

**MATERIALS AND METHODS**

**Enzymes**

For PCR, the following thermostable prokaryotic polymerases were used; *Taq*, *ExTaq* (TAKARA, Tokyo, Japan) and *Pfu* (Stratagene, La Jolla, CA, USA). T4 DNA polymerase was purchased from TAKARA Co. Ltd (Tokyo, Japan).

The conditions for reaction with each polymerase are as follows: *Taq*, 10 mM Tris–Cl pH 8.0, 1.5 mM MgCl₂, 50 mM KCl; *Pfu*, 20 mM Tris–HCl pH 8.0, 2.0 mM MgCl₂, 10 mM KCl, 10 mM (NH₄)₂SO₄, 2 mM MgSO₄, 0.1% Triton X-100, 0.1 mg/ml bovine serum albumin (BSA).

**Cell lines and tissue specimens**

Human colon adenocarcinoma cell lines, HCT116, LS174T and SW48, were obtained from American Type Cell Culture (ATCC) and maintained in suitable media supplemented with 10% fetal calf serum. Single colony isolation was performed using 96-well plates. Each colony was confirmed microscopically and isolated by trypsinization.

All the cancer and normal tissue specimens were obtained at the time of surgical treatment of cancer patients in the Department of Surgery II, Kyushu University Hospital.

**Oligonucleotides used for PCR**

The 105 bp sequence of D13S175 microsatellite marker including 17 copies of (CA) repeat was artificially synthesized and purified using high performance liquid chromatography (HPLC) and polyacrylamide gel electrophoresis (PAGE). This artificially synthesized microsatellite DNA was used as a template for PCR. Other oligonucleotides were synthesized and purified by HPLC. The sequences of the oligonucleotide primers for PCR are as follows: D2S123-5′, 5′-AAACAGGATGCCCTCCTTTA; D2S123-3′, 5′-GGACCTTCCACCTATGGGAC; D5S107-5′, 5′-GGCATCAACTTGAAACAGCAT; D5S107-3′, 5′-GTTCACCTTAAACAATAC; D10S197-5′, 5′-ACCAGCTCAGTGGTACGAC; D10S197-3′, 5′-GTGATCTGTCCTCAGTTGCC; D11S904-5′, 5′-ATGACAACAACTTTCAGGAC; D11S904-3′, 5′-CTGTGTTATATCCCTAAGTGTA; D13S175-5′, 5′-TGCACTCAGTATAGAATTTAGTTA; D13S175-3′, 5′-TATTGATATCTGGAATCTGGCT.

The 5′ PCR primers were labeled with ROX (6-carboxy-x-rhodamine) or HEX (6-carboxy-2′,4′,7′,4,7-hexachloro-fluorescein). Size markers were labeled with TAMRA (N,N,N′,N′-tetramethyl-6-carboxy-rhodamine).

**Preparation of genomic DNA**

Genomic DNA extracted from cultured cells and surgical specimens was also used as a template for PCR. Tissue was ground in liquid nitrogen. Cells were lysed in digestion buffer (10 mM Tris–Cl pH 8.0, 0.1 M EDTA pH 8.0, 0.5% SDS, 20 µg/ml pancreatic RNase). After treatment of proteinase K and extraction with phenol, DNA was precipitated with ethanol, then dissolved in 1× TE (10 mM Tris–Cl pH 7.5, 1 mM EDTA). The concentration of DNA was determined by OD260 using a spectrophotometer.

**PCR using synthesized templates**

The PCR reactions using synthesized oligonucleotide templates were performed using TAKARA *Taq* Reagent Kits (TAKARA Co. Ltd, Tokyo, Japan) and run in the Perkin Elmer GeneAmp PCR system 9600 or 2400 (Norwalk, CT, USA). A 50 µl reaction mixture contained 1× reaction buffer, 350 µM each dNTP, 10 pmol each primer, 2.5 U of the polymerase and 1.0 fg template DNA. The thermal conditions of the system were as follows: one cycle at 95°C for 4 min; 35 cycles at 95°C for 30 s, 55°C for 30 s and 72°C for 30 s; one cycle at 72°C for 10 min. Then 0.5 U T4 DNA polymerase was added to the mixture and incubated at 37°C for 10 min. An aliquot (1.5 µl) of the product was mixed with 0.5 µl loading buffer (blue dextran, 25 mM EDTA), 2.5 µl formamide and 0.5 µl TAMRA-labeled size marker, and then denatured and loaded onto the ABI 373A sequencer. To compare two samples, 1.2 µl of ROX-labeled product and 0.3 µl HEX-labeled product were mixed and loaded, as described above. A size marker was always electrophoresed in each lane to standardize the mobility of the sample. The running condition was 1500 V, 20 mA, 30 W, for 5.5 h. The data was processed using ABI software, GeneScan.

**PCR using genomic DNA**

The PCR using genomic DNA templates was performed as described above with the exception that the reaction mixture contained 25 ng genomic DNA as a template. DNA derived from samples was amplified with ROX-labeled 5′ primer and cold 3′ primer and DNA from controls with HEX-labeled 5′ primer and cold 3′ primer. Both PCR products were electrophoresed in the same lane, using an ABI 373A sequencer.

**RESULTS**

Optimization of electrophoretic profiles of amplified microsatellite sequences using artificial templates

To optimize the PCR conditions for microsatellite analysis, we used artificially synthesized microsatellite alleles. As shown in Figure 1A, by PCR the 105 bp region of an artificially synthesized D13S175 microsatellite sequence was ampliﬁed as a complex of some fragments of different lengths. The peak corresponded to 108 bp and was accompanied by shorter bands different by one base. Since heterogeneity of templates was excluded, this multiplicity of bands must be derived from slippage and terminal deoxynucleotidyl transferase activity of *Taq* polymerase. In PCR, other thermostable prokaryotic DNA polymerases are available and these often have 3′–5′ exonuclease activity. We next used *Pfu* and *Taq* conferred exonuclease activity (ExTaq) to improve the electrophoretic proﬁle of the amplified microsatellite sequences. However, the scanned bands were essentially the same as observed with use of *Taq* polymerase (data not shown). On the other hand, after treatment with T4 DNA polymerase which also has 3′–5′ exonuclease activity, the amplified products consisted of only 107, 105 and 103 bp bands (Fig. 1B). These findings mean
Figure 1. Amplification of artificially synthesized microsatellite alleles with fluorescent PCR primers. (A) An artificially synthesized D13S175 microsatellite allele amplified with \textit{Taq} polymerase; single-stranded 105 bp DNA corresponding D13S175 microsatellite marker was artificially synthesized and 1.0 fg of the oligonucleotide was amplified by PCR with \textit{Taq} polymerase. The amplified product was mixed with loading buffer, denatured and electrophoresed in the ABI 373A sequencer. The result in fluorescence scanning was processed by the ABI software, GeneScan. (B) \textit{T4} DNA polymerase treatment of microsatellite sequences amplified from the artificial D13S175 allele. The D13S175 sequences were amplified from the artificially synthesized templates with \textit{Taq} polymerase and treated with 0.5 U \textit{T4} DNA polymerase at 37°C for 10 min and then electrophoresed in an ABI 373A sequencer.

That with slippage of \textit{Taq} polymerase, there is one added or subtracted copy of (CA) repeat in the amplified products of D13S175 sequence and part of each fragment harbors one additional base at the 3' terminus of the fragment. As shown in Figure 1B, the most simplified profile of amplified microsatellite sequences is a two bases-pitched and single-peaked cluster of bands showing an increasing pattern; the amount of fragments increases in proportion to their length. When using the genomic template, the electrophoretic profiles of amplified microsatellite sequences were clearly simplified in case of \textit{T4} DNA polymerase treatment (Fig. 2). Even when profile of the amplified products was two bases-pitched originally, it became a two bases-pitched cluster of bands shorter by one base than before treatment (Fig. 2, left). Thus, the complexity of amplified microsatellite sequences observed in genomic analysis is not due to heterogeneity of templates but rather to the intrinsic property of \textit{Taq} polymerase and that the complexity can be remarkably reduced by enzymatic modification of the amplified products.

Dual fluorescence co-electrophoresis for microsatellite instability analysis

To verify the detection characteristic of our system, we next examined the signal magnitude in case of various amounts of PCR product. The detection characteristic was completely linear in the observed range but the specific intensities differed in labeling compounds (Fig. 3A). Therefore, in our system, bands are detected quantitatively in a wide signal range (Fig. 3B).

In microsatellite instability analysis one should compare the results derived from two independently amplified products in the same lane in order to minimize migration errors. However, the electrophoretic mobilities are different in fragments labeled with each fluorescence compound. Since the electrophoretic mobilities of ROX and HEX proved to be the closest (data not shown), we chose these two labeling compounds for the system. To compare the results from two independently amplified products within a similar signal magnitude, we set the ratio of mixture of products amplified with ROX- and HEX-primers in 4:1, as deduced from the data in Figure 3A. Therefore, in any cases negative in microsatellite instability, the electrophoretic profiles

Figure 2. \textit{T4} DNA polymerase treatment of microsatellite sequences amplified from genomic DNA. The D2S123 and D13S175 sequences were amplified from genomic templates. The amplified products were treated with 0.5 U \textit{T4} DNA polymerase at 37°C for 10 min, and then electrophoresed in an ABI 373A sequencer. The results before and after treatment were superimposed in a display.
Figure 3. Detection characteristics in fluorescent microsatellite analysis. (A) Various amounts of PCR products amplified with a fluorescence-labeled oligonucleotide primer were loaded onto an ABI 373A sequencer. The results of scanning in two fluorescence labeling, ROX and HEX, are shown. (B) The DSS107 microsatellite sequence was independently amplified with HEX-primer and ROX-primer and co-electrophoresed in the same lane. In the lower column, the signal magnitudes of the both amplified products were adjusted by calculation in the software, GeneScan.

Figure 4. Detection sensitivity of heterogeneity of repeat number in mixtures of microsatellite templates. In the subclones of HCT116 cells, we selected two independent clones harboring DSS107 sequences with (CA) repeats differing only by two bases. Genomic templates derived from these two clones were mixed at various ratios, then the mixtures were amplified as templates by PCR and analyzed.

of two independently amplified products match completely in our system (Fig. 3B).

Sensitivity of the high resolution fluorescent microsatellite analysis

We next assessed the sensitivity of our system, which we call ‘high resolution fluorescent microsatellite analysis’, using a mixture of templates including different numbers of dinucleotide repeats. In independently replicating clones of colon carcinoma cell line, HCT116, which has a mutation in the hMLH1 gene and is mismatch repair-defective, we found that the numbers of dinucleotide repeats in certain microsatellite sequences varied widely (unpublished data). In these subclones of HCT116 cells we found two independent clones harboring DSS107 sequences with (CA) repeat stretches different only by two bases. Using mixtures of genomic templates derived from these two clones we examined the sensitivity to detect templates different in repeat number. We first mixed two genomic templates at various ratios, then, we amplified each template mixture by PCR and analyzed the products. The system readily revealed the existence of (CA)$_{n+1}$ templates in the 1:4 mixture (Fig. 4), the limit of detection was 10%. On the other hand, the limit of detection of (CA)$_{n-1}$ templates may be 50%, since the existence of (CA)$_{n}$ templates in (CA)$_{n+1}$ templates was scarcely detected in the 1:1 mixture. Thus, we demonstrated that fluorescent microsatellite analysis is sufficiently sensitive to heterogeneity of dinucleotide repeat number in given DNA templates, especially in cases of detection of templates with excessive copies of dinucleotide repeats.

Classification of microsatellite alteration

Using the system developed here, we analyzed microsatellite alterations in several cancer cell lines and various human cancers. The system efficiently detected diverse types of microsatellite alterations including extremely subtle ones (data not shown). Through these analyses, we found that all the profiles of amplified microsatellite sequences derived from summation of basic profiles; a two bases-pitched and single-peak cluster of bands showing an increasing pattern. From this point of view, we classified all the patterns of microsatellite alteration to deduce new criteria for microsatellite instability (Fig. 5). We first divided all the patterns of alteration into two groups: Group A, when the clusters of bands derived from paternal and maternal alleles cross in scanning; Group B, when the clusters of bands derived from the both alleles are separated in scanning. Groups A and B consist of four and two subgroups, respectively. Basic alterations implying microsatellite instability are (i) appearance or loss of bands and (ii) shift of the peak. Practical examples of alteration categorized in each subgroup are shown in Figure 6.

DISCUSSION

In the new system we have developed electrophoretic profiles of PCR-amplified microsatellite sequences are markedly improved and optimized for judgment of microsatellite instability. Using artificially synthesized templates we found that the multiplicity of bands in scanning were not due to heterogeneity of genomic templates but rather to the intrinsic property of Taq polymerase, slippage and the terminal deoxynucleotidyl transferase activity of the enzyme. The latter is to add an extra one base, usually adenine,
Figure 5. Classification of microsatellite alteration in high resolution fluorescent microsatellite analysis. Classifying the practical examples, all the patterns implying microsatellite instability are illustrated. Dotted lines indicate patterns of normal controls and solid lines indicate patterns derived from samples.

at the 3' end of the synthesized strands in a template-dependent manner (12). This activity of Taq polymerase is variably expressed in a sequence-dependent manner (12,13) and consequently alteration by 1 bp of amplified products, referred to as 'shadow bands' or 'stuttering' (14), is not reproducible. Therefore, these additional bands should be removed for a precise assessment of microsatellite sequences. As Ginot et al. pointed out (15), only T4 DNA polymerase could eliminate these additional bases from the amplified fragments. Although efforts have been made to simplify the electrophoretic profile by modifying PCR primers (13), T4 DNA polymerase treatment may be the simplest. Thus, when one uses PCR for microsatellite analysis, clusters of several bands of different lengths have to be observed. However, 'shadow bands' or 'stuttering' should be erased for a precise assessment. The finally optimized profile of amplified microsatellite sequences is a two bases-pitched cluster of bands, which is a summation of the basic profile; a two bases-pitched, single-peaked cluster of bands showing an increasing pattern.

Migration errors were ruled out in our system. With conventional methods, two amplified products derived from normal tissues and tumors are independently electrophoresed in separate lanes. However, these methods are susceptible to migration errors, which means a misread of changes in band position. Thus far, some fluorescent systems have been described in which two amplified products derived from normal tissues and tumors are independently electrophoresed in separate lanes and the results are superimposed in calculation (10,11). However, even this method cannot completely exclude migration errors. We examined electrophoretic mobilities and specific intensity of fragments labeled with various fluorescence compounds and found the combination in which the phases of electrophoresis profiles matches completely. Thus, this is the first report of a dual fluorescence co-electrophoresis system in which migration errors have been completely excluded.

Detection errors can be minimized with our system. We demonstrated that our fluorescent microsatellite analysis has consistently linear detection characteristics, while those of autoradiography are sigmoidal. As changes in number or height
of the bands are extremely important in judgment of microsatellite instability, widely linear characteristics are to be desired. We also determined the ratio of magnitude to adjust the signal magnitudes of two amplified products labeled with different fluorescences. Efficiency of amplification can differ in two independent PCRs and signal magnitudes of amplified products will consequently differ. In such cases, one can adjust them by calculation in the software, GeneScan, since the detection characteristics are linear in the range. However, signal magnitudes of two samples labeled with the same fluorescent compound cannot be adjusted in GeneScan. Here again, the advantage of the dual fluorescence co-electrophoresis system was made evident.

Thus, we have established a system to compare two electrophoretic profiles of microsatellite sequences within the same phase. Using genomic DNA mixtures including different copy numbers of (CA) repeat, \((\text{CA})_{n+1}\) or \((\text{CA})_{n-1}\), in D5S107 microsatellite locus, we evaluated sensitivity of the system to detect the heterogeneity of (CA) repeat numbers in a given template DNA. This evaluation seems important, since tumor tissues are often heterogeneous in genotype and include normal cells. In the system developed here, one can detect the existence of templates harboring an excessive (CA) repeat in the 1:9 mixture. Such high sensitivity is attributed to widely linear detection characteristics of this system and is sufficient to assess microsatellite instability in clinically obtained specimens.

In addition to methodological problems of the conventional microsatellite analyses, the criteria for judgment of microsatellite instability have been diverse. In the conventional method using autoradiography, Thibodeau et al. (8) based judgment on the appearance of new bands differing by more than two bases. Analyzing many cell lines and tissue specimens using high resolution fluorescent microsatellite analysis, we classified all the patterns judged to be positive in microsatellite instability (Fig. 5) and deduced new criteria for determining microsatellite instability. Since independently replicating clones of human normal fibroblast cell lines were reported to show no alteration of microsatellite sequences (16), we regard a given tumor as positive in microsatellite instability when the alterations listed in Figure 5 are obvious in at least one locus of five independent microsatellite loci, or when the alterations are suspicious in more than one of five loci. This criteria is applicable under the condition that any ‘shadow bands’ or ‘stutterings’ are completely eliminated. Under the present conditions some positive patterns cannot be distinguished from loss of heterozygosity (LOH), yet some investigators assert it is possible to distinguish them (9).

The rate of positivity for microsatellite instability was reported to reach 90% in patients who had more than one G.I. tract malignancy and 30% in patients of sporadic gastrointestinal cancer (6). These high rates of positivity suggest the clinical significance of microsatellite instability in assessment of the risk of new or secondary malignancies. Using the high resolution fluorescent microsatellite analysis developed here and the criteria we proposed, a precise judgment can be made and rates of positivity in various human malignancies might be corrected.

ACKNOWLEDGEMENTS

We thank K. Takahashi, M. Sakabe and Y. Ohkuma (Perkin Elmer Japan Co., Ltd Applied Biosystems Division) for technical support. We are most grateful to M. Sekiguchi for helpful advice and to M. Ohara for comments on the manuscript. This study was supported by grants from the Ministry of Education, Science, Sports and Culture of Japan.

REFERENCES