Rapid and accurate approach for screening of microsatellite unstable tumours using quasimonomorphic mononucleotide repeats and denaturating high performance liquid chromatography (DHPLC)

Gašper Berginc and Damjan Glavac

Department of Molecular Genetics, Institute of Pathology, Faculty of Medicine, University of Ljubljana, Korytkova 2, 1000 Ljubljana, Slovenia

Abstract. MSI analysis is becoming increasingly important for the detection of both hereditary non-polyposis colorectal cancer and sporadic primary colorectal tumours with MSI high phenotype. The Bethesda panel of five microsatellite markers has been proposed to provide uniform criteria for MSI analysis. Here we report on an MSI analysis approach using quasimonomorphic mononucleotide repeats and denaturating high performance liquid chromatography (DHPLC). We analysed 595 newly diagnosed colorectal tumours and 145 normal samples. Microsatellite markers BAT-25, BAT-26, NR-21, NR-22, and NR-27 were amplified in multiplex reaction and analysed using DHPLC and capillary electrophoresis (CE). DHPLC conditions for analysis of MSI multiplex assay were evaluated and tested. Analysis and cross-examination of the results obtained from 96 samples using DHPLC and capillary electrophoresis showed the same sensitivity and specificity of the two approaches for detecting MSI-H tumours. Using our new approach we showed that the tested markers are quasimonomorphic in a Slovenian population, with frequencies of polymorphisms 0.07%, 1.4%, 2.1%, 1.4%, and 1.4% for BAT-25, BAT-26, NR-21, NR-22, and NR-27, respectively. Forty-three (7.2%) new MSI-H tumours were identified, of which 84% showed instability in all 5 tested markers. Overall, we developed a high-throughput, robust, accurate and cost-effective approach for the detection of MSI-H tumours.

Keywords: Microsatellite instability, quasimonomorphic mononucleotide repeats, multiplex PCR, DHPLC

1. Introduction

Microsatellite instability (MSI) is a phenomenon first described in colorectal cancer (CRC) and characterized by small deletions or insertions within short tandem repeats in tumour DNA compared to matching normal DNA. Approximately 10 to 15 percent of sporadic CRC and the majority of hereditary non-polyposis colorectal cancers show MSI [1–4]. MSI is also seen in a significant proportion of other extracolonic cancers [5]. The Bethesda panel of two mononucleotide (BAT-25, BAT-26) and three dinucleotide (D5S346, D2S123, D17S250) microsatellite markers was proposed by the National Cancer Institute Workshop in 1997. In terms of MSI status, tumours can be classified into three groups. The first is characterized as MSI-high (MSI-H, showing instability in ≥40% of tested markers); the second as MSI-low (MSI-L, showing instability in <40% of tested markers); and the third shows no instability (MSS, microsatellite stable). MSI-
H is a distinct group of tumours associated with specific clinicopathological features and a favourable prognosis [6]. The distinction between the MSI-L and MSS groups is still a matter of debate, due to contradictory data [7–11]. The determination of MSI status has become a very important tool in hereditary non-polyposis colorectal cancer (HNPCC) screening and predicting the responsiveness to chemotherapy [12,13].

A lot has been done in MSI analysis in recent years concerning accuracy and speed. Developments in PCR technology have enabled co-amplification of several microsatellite markers in a single multiplex reaction [14–16]. In order to determine the best set of microsatellite markers for the identification of MSI-H tumours and to simplify MSI analysis, extensive studies have been done in relation to sensitivity and specificity. The Bethesda panel provides a uniform set of markers and criteria in MSI analysis, although it has some limitations because it uses three dinucleotide markers in the panel [17]. The dinucleotide repeats in this panel generally show instability in only 60–80% of MSI-H tumours [14]. Their highly polymorphic nature requires the analysis of tumour and corresponding germline DNA (not always available), and it results in the misclassification of MSI if samples from two individuals are mixed [18]. Misinterpretation of allelic profiles using dinucleotide repeats can also be a consequence of mutations in the hMSH6 gene, causing primary alterations in mononucleotide repeats [19], and the presence of stutters caused by polymerase chain reaction (PCR) artefacts produced by DNA polymerase slippage [20,21]. To overcome these problems, several authors have suggested the use of BAT-26 alone for the evaluation of MSI, without the need for corresponding germline DNA analysis, due to its quasimonomorphic nature [22–24]. Population-based studies have shown that most Caucasian populations have less than 1% of polymorphic alleles on the BAT-26 locus but among African-American populations it is polymorphic in 7–13% of alleles [25,26]. These data suggest that misclassification of MSI can occur in different populations if only BAT-26 is analysed without matched normal tissue DNA. In an effort to improve the existing panel of MSI markers proposed by the Bethesda guidelines, more mononucleotide markers have been tested for germline polymorphisms and several of them have shown a quasimonomorphic nature [20,23]. Suraweera et al. [14] proposed a set of five quasimonomorphic mononucleotide microsatellite markers (BAT-25, BAT-26, NR-21, NR-22, and NR-24) analysed together in pentaplex PCR, with nearly 100% sensitivity and specificity, eliminating the need for analysis of corresponding germline DNA. A tumour is defined as MSI-H when at least three out of five mononucleotides show instability [14]. With polymorphisms occurring in 1% of the Caucasian population and 10% of the African-American the probability of having 3 polymorphic markers would be $10^{-6}$ and $10^{-3}$ respectively [15].

Conventional methods used for MSI analysis are non-automated single-stranded conformation analysis (SSCA) and semi-automated fluorescence-based capillary electrophoresis (CE). Potocnik et al. [27] identified 10% of MSI-H tumours in 345 tested samples using SSCA. In their studies, Aaltonen et al. [28] identified 12% of MSI-H tumours (509 samples tested), Pinol et al. [29] identified 6.8% of MSI-H tumours (1222 samples tested), and Salovaara et al. [30] identified 12% of MSI-H tumours (535 samples tested) with fluorescence-based CE. Several weaknesses were observed using these methods. The first method is labour-intensive, time consuming, and needs strict electrophoretic conditions; the second needs fluorescently labelled primers and additional software [31]. The aforementioned drawbacks cause lower throughput and higher costs of MSI detection, making it unsuitable for large scale and everyday screening.

The purpose of our study was to develop a fast, accurate, and cost-effective approach for detecting MSI-H tumours. Denaturing high performance liquid chromatography (DHPLC) is a fully automated, high-throughput method with various applications. It has primarily been used for the detection of mutations and single nucleotide polymorphisms under partially denaturating conditions but various other applications are possible (dsDNA sizing, gene mapping, gene expression analysis, methylation analysis and others). The main advantages of DHPLC are that it does not require modified PCR primers, it does not use specific reagent arrays, it does not have specific detection labels, and it does not require additional sample treatment [32]. Using a multiplex system with a set of quasimonomorphic mononucleotide markers [14,15,20] and DHPLC we managed to minimize the time and costs of analysis without reducing the accuracy. Using our new approach, we screened Slovenian colorectal cancer patients, identified 43 new MSI-H patients, and checked for polymorphisms in mononucleotide microsatellite markers.
2. Materials and methods

2.1. Samples

We analysed 595 colorectal tumours from all over Slovenia, newly diagnosed between 1999 and 2004, and 145 normal DNA samples. All samples were gathered by the Institute of Oncology in Ljubljana, where snap frozen biopsies of colorectal lesions were histologically evaluated. Primary colorectal adenocarcinomas and corresponding normal tissue taken from a site distant to the tumour were used in the study. For evaluation of the specificity and sensitivity of our new approach, we used 17 MSI-H samples from a previous study [27]. The MSI status of these samples was determined using a panel of microsatellite markers proposed by the Bethesda guidelines [6]. The Medical Ethical Commission of Republic of Slovenia approved this study in April 2006.

2.2. DNA isolation

DNA was extracted from snap frozen colorectal tumours and corresponding normal tissue using NucPrep™ chemistry (Applied Biosystems, Foster City, CA, USA) on an ABI 6100 (Applied Biosystems, Foster City, CA, USA).

2.3. Primers and PCR

Different sets of five quasimonomorphic mononucleotide microsatellite markers BAT-25, BAT-26, NR-24, NR-21, NR-27 were amplified in multiplex PCR reactions using previously described primers and PCR conditions [15] with HotMaster™ DNA Taq polymerase (Eppendorf AG, Hamburg, Germany). We used two sets of primers; one set with a fluorescently labelled anti-sense primer for analysis on an ABI 310 (Applied Biosystems, Foster City, CA, USA) genetic analyser and one set without fluorescent dyes for DHPLC analysis on the WAVE system (Transgenomic Inc., Omaha, NE, USA).

2.4. Microsatellite instability analysis by capillary electrophoresis

Microsatellite instability analysis was performed simultaneously for all five amplified microsatellite markers on an ABI 310 (Applied Biosystems, Foster City, CA, USA) automated genetic analyser. GeneScan 3.7 software (Applied Biosystems, Foster City, CA, USA) was used for data analysis. Samples showing instability in 3 or more markers were classified as MSI-high.

2.5. Microsatellite instability analysis by DHPLC

DHPLC analysis was performed on a fully automated WAVE system (Transgenomic Inc., Omaha, NE, USA). PCR products were injected into a DNASEp cartridge (Transgenomic Inc., Omaha, NE, USA) and eluted at a flow rate of 0.9 ml/min through a linear gradient of acetonitrile containing 0.1 M triethylammonium acetate (TEAA). The elution gradient of Buffer A (0.1 M TEAA solution) and Buffer B (0,1 M TEAA containing 25% acetonitrile solution) was automatically adjusted using WAVemaker™ software (Transgenomic Inc., Omaha, NE, USA) (Table 1). Oven temperature was set at 50°C as recommended for double-stranded DNA sizing. A run was performed using a multiple fragments double-stranded DNA sizing application which allowed us the simultaneous analysis of microsatellites. Gradient conditions were set for the elution of 100 base-pair products in 4.5 minutes. The sizes of the largest and smallest PCR products were set at 77 bp and 189 bp, respectively. UV detection was performed at a 260 nm wavelength.

3. Results

In order to discover whether using DHPLC enabled us to identify all 5 microsatellite markers in the MSI multiplex assay proposed by Buhard et al. [15], we first tested 145 normal samples. Following several optimisation steps, we managed clearly to see all the microsatellite markers multiplied in a single pentaplex PCR (Fig. 1B). The same normal samples were analysed for the presence of polymorphisms in the Slovenian population. All 5 microsatellite markers showed a quasimonomorphic nature, with percentage of polymorphisms not exceeding 2.1 (BAT-26 0.07%, BAT-25 1.4%, NR-24 2.1%, NR-27 1.4%, and NR-27 1.4%).

We tested the MSI multiplex assay on 17 previously identified MSI-H samples. Comparing the MSS and MSI-H DHPLC elution profiles, there was an obvious change, with additional peaks appearing in the MSI-H sample (Fig. 1D). There were no differences between the previously evaluated CE-based method (Fig. 1C) [15] and our new DHPLC-based approach in identifying the MSI-H samples; both approaches identified all MSI-H samples. Using both the aforementioned approaches, we analysed and cross-examined 96 newly diagnosed colorectal cancers and identified 8 new MSI-H samples. The results of the two approaches were 100% identical in MSI-H identification. A
comparison of the MSI status of all markers tested in 96 samples revealed that our new DHPLC-based approach correctly identified 98% (472/480) when using the CE-based method as a reference. Misclassification of some markers occurred due to the complex pattern of novel and shifted allele peaks (Fig. 2A), but none of the misclassified markers affected the identification of MSI-H samples. In order to solve the problem of misclassification, we amplified 5 microsatellite markers in two separate multiplex PCR reactions, one including three (NR-27, NR-24, BAT-26) and the other including two (NR-21, BAT-25) markers. When analysed by DHPLC, the elution profiles of triplex and duplex PCR reactions showed a simplified pattern of allele peaks, due to a bigger size difference among the PCR products of the analysed microsatellite markers, which enabled us unambiguously to assign novel allele peaks to each individual marker (Fig. 2B and 2C). The DHPLC-based approach with a reduced number of markers per analysis was able to detect short deletions in individual markers (Fig. 2B). MSI results for each individual marker obtained on 96 cross-examined tumours by DHPLC analysis of two separate PCR reactions (triplex, duplex) were 100% identical to the results obtained by CE.

Because of the high specificity and sensitivity of our new approach, we analysed an additional 499 newly diagnosed colorectal cancers. Overall, we tested 595 newly diagnosed colorectal cancers and identified 43 (7.2%) new MSI-H tumours. Among the 43 MSI-H tumours, 36 (84%) were unstable at all five markers and 7 (16%) were unstable at four markers.

4. Discussion

Both large-scale studies and everyday diagnostic procedures need fast, accurate, and cost-effective ap-

Table 1

| Time (min) | Buffer A (%) | Buffer B (%) | Oven temp. (°C) | Flow rate (ml/min) |
|------------|--------------|--------------|-----------------|-------------------|
| 0.0        | 58           | 42           | 50              | 0.9               |
| 0.5        | 53           | 47           | 50              | 0.9               |
| 1.8        | 49           | 51           | 50              | 0.9               |
| 3.0        | 47           | 53           | 50              | 0.9               |
| 4.3        | 44           | 56           | 50              | 0.9               |
| 5.5        | 43           | 57           | 50              | 0.9               |

Fig. 1. MSI analysis using capillary electrophoresis and DHPLC. (A) typical electropherogram of an MSS tumour. (B) typical DHPLC elution profile of MSS tumour. (C) typical electropherogram of a MSI-H tumour. (D) typical DHPLC elution profiles of MSI-H tumours compared with elution profile of MSS tumour. Arrows indicate the presence of new sized alleles.
Fig. 2. Comparison of DHPLC elution profiles and electropherogram of a sample in which misclassification of the MSI status occurred at the NR-24 marker due to a complex elution profile and a short deletion. (A) complex DHPLC elution profile of MSI-H sample analysed as a single pentaplex reaction. (B) DHPLC elution profile of MSI-H sample, in which 3 mononucleotide markers are analysed simultaneously. (C) DHPLC elution profile of MSI-H sample in which 2 mononucleotide markers are analysed simultaneously. (D) CE electropherogram of MSI-H sample. Arrows indicate new sized alleles.

Table 2

|                      | CE (5 markers) | DHPLC (5 markers) | DHPLC (3 markers) |
|----------------------|---------------|-------------------|-------------------|
| Accuracy<sup>1</sup>:|               |                   |                   |
| – detection of MSI-H tumours | reference     | 98%               | 100%              |
| – detection of MSI at specific marker | reference     | 100%              | 100%              |
| Cost of analysis<sup>2</sup> | 2.14 EUR/sample | 1.92 EUR/sample | 1.90 EUR/sample |
| Cost of apparatus<sup>3</sup> | ~ 55,000 EUR   | ~ 100,000 EUR    | ~ 100,000 EUR    |
| No. of steps in analysis<sup>3</sup> | 5             | 2                 |                   |
| Rapidity of the method<sup>3</sup> | ~ 30 min/sample | ~ 9 min/sample   |                   |

<sup>1</sup>Expressed as the concordance of the results obtained by the new DHPLC-based approach with the results obtained by the previously evaluated CE-based approach (reference).

<sup>2</sup>Approximate price includes the costs of apparatus and software, but the prices are not comparable since DHPLC is sold in a bundle with more complex software.

<sup>3</sup>In addition to PCR amplification and sample loading, 3 additional steps are needed in CE analysis (sample dilution, mixing of sample with size standard and denaturation).

proaches for MSI analysis. Methods conventionally used for MSI analysis (SSCA, fluorescence-based capillary electrophoresis) do not fulfil all of the proposed criteria [31,33]. Gel-based SSCA analysis is widely used because it is inexpensive and does not need additional equipment. However, SSCA does not meet the criteria for a high-throughput method because of its several labour-intensive steps, including making the gel, loading samples, lengthy staining procedures, and difficulties in data interpretation [31]. CE solved several of the aforementioned problems, but there are still several drawbacks [33]. The need for specific reagents (fluorescently labelled primers, size standard) and additional steps in preparation of the sample for analysis affect the cost-effectiveness and labour-intensiveness of this method (Table 2). Alternative methods using DHPLC have been proposed for MSI analysis [31,33]. Two different approaches have been described; the first used non-denaturing conditions [31] and the second used fully denaturing conditions [33]. Simple elution
profiles with one peak per marker under non-denaturing conditions allowed us to perform multiplex analysis, which would be very difficult under fully denatured conditions as described by Pan et al. [33], due to several peaks appearing for each tested marker and possible overlapping in multiplex analysis. In order to develop a new approach for MSI analysis using DHPLC technology, we chose the quasimonomorphic mononucleotide microsatellite markers (BAT-25, BAT-26, NR-24, NR-21, and NR-27) proposed by Buhard et al. [15]. We tested 145 normal DNA samples in order to prove the quasimonomorphic nature of these markers in a Slovenian population. The observed percentages of polymorphisms (0.7–2.1%) are in accordance with the percentages described for Caucasian populations [14,25,26]. The high sensitivity and specificity of the aforementioned microsatellite markers in combination with their quasimonomorphic nature allowed us to omit the analysis of germline DNA corresponding to tumour [14,15]. This enabled us radically to reduce the time needed for the analysis of a single sample, without decreasing accuracy. Our approach enabled us to identify the MSI status of a single sample in as little as 9 minutes, which is one third of the time needed for MSI analysis using capillary electrophoresis (Table 2).

Several optimisation steps were necessary in order clearly to see all markers in the DHPLC elution profile. The PCR product of NR-27 (length of 87 base pairs) was shorter than the 100 base pairs recommended by the manufacturer. We observed that buffer gradient conditions calculated by W A VEMAKER software for multiple fragment dsDNA sizing analysis, with the exact length of the shortest fragment inserted, did not work out well (Fig. 3). Under these conditions we could not see the shortest fragment (NR-27) because it was eluted from the column simultaneously with the PCR primers and hidden in the first peak of the elution profile. We overcame this problem by inserting a value 10 bp lower than the real length of the shortest PCR product. All peaks were thereafter seen clearly on the elution profile, as shown in Fig. 1B. The interpretation of mononucleotide microsatellite markers is a straightforward procedure due to their single peak appearance on an elution profile under non-denaturing conditions. Any additional peak on the elution profile represents an allele of a different size, which indicates MSI in a given marker. When MSI-H samples were analysed using all 5 markers amplified in a single pentaplex PCR, a complex pattern of allele peaks, in combination with very short deletions at some markers, caused ambiguity in the detection of MSI in a specific allele (Fig. 1D and 2A). These problems resulted in misclassification of 2% of all tested markers among 96 samples cross-examined by CE. In order to obtain simplified elution profiles, we reduced the number of simultaneously analysed markers. With triplex and a duplex PCR reactions analysed separately, we obtained clear elution profiles, which enabled us clearly to identify all novel peaks and assign them to specific markers (Fig. 2B and 2C). The MSI results of each individual marker obtained by the DHPLC-based approach with a reduced number of simultaneously tested markers, and the CE-based approach, were 100% identical. Because all the tested microsatellite markers are very specific for the detection of MSI, the elimination of two markers does not affect the overall sensitivity and specificity of the analysis. According to Bethesda guidelines (at least 40% of positive tested markers) 2 out of 3 markers should detect MSI in order to classify the tumour as MSI-H. Because of the percentages of detected polymorphisms in the Slovenian population, the chance of there being 2 polymorphic markers when using a combination of NR-27, NR-24 and BAT-26, is between 3$^{-4}$ and 10$^{-6}$, which allowed us to eliminate the analysis of germline DNA, even though the number of analysed markers was reduced. When only one marker is positive, 2 additional markers (BAT-25 and NR-21) should be tested and the guidelines for pentaplex MSI assay described by Buhard et al. [15] should be followed for MSI classification. There were no cases in our study in which the analysis of 2 additional markers would affect the MSI-H classification, because all detected MSI-H samples showed instability in at least 4 out of 5 microsatellite markers. Of 597 tested samples, only 7 (1.2%) showed instability at 1 of the 3 tested markers and therefore the need to test an additional 2 markers. According to our results, the analysis of 3 mononucleotide microsatellite markers (NR-27, NR-24, BAT-26) by the newly developed DHPLC-based approach showed the same specificity and sensitivity as the CE-based approach. It is therefore a feasible alternative, and has the advantage of its high-throughput and reduced costs (~11%).

A total of 43 (7.2%) out of 595 colorectal cancer patients were found to have mismatch repair deficiency. The slightly lower percentage of identified MSI-H tumours than previously published (10–15%) [27, 28,30] might be due to previous overestimations because of the use of polymorphic dinucleotide markers (2 positive dinucleotide markers may not detect true MSI-H tumours), potential errors due to mixing of germline and tumour DNA, and differences between populations [15].
In summary, we developed a high-throughput, robust, accurate and cost-effective approach for detecting MSI-H tumours using DHPLC. The approach is suitable for large-scale studies as well as for everyday use with smaller numbers of samples and performs well even in populations with a higher percentage of polymorphic microsatellite markers. Use of quasi-monomorphic mononucleotide markers on fully automated DHPLC enabled us to obtain accurate information about the MSI status of a tumour in only nine minutes and to reduce the price of analysis.

References

[1] Y. Ionov, M.A. Peinado, S. Malkhosyan, D. Shibata and M. Perucho, Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis, *Nature* 363 (1993), 558–561.

[2] L.A. Aaltonen, P. Peltomaki, F.S. Leach, P. Sistonen, L. Pylkkanen, J.P. Mecklin, H. Jarvinen, S.M. Powell, J. Jen and S.R. Hamilton, Clues to the pathogenesis of familial colorectal cancer, *Science* 260 (1993), 816–819.

[3] S.N. Thibodeau, G. Bren and D. Schaid, Microsatellite instability in cancer of the proximal colon, *Science* 260 (1993), 816–819.

[4] P. Peltomaki, L.A. Aaltonen, P. Sistonen, L. Pylkkanen, J.P. Mecklin, H. Jarvinen, J.S. Green, J.R. Jass, J.L. Weber and F.S. Leach, Genetic mapping of a locus predisposing to human colorectal cancer, *Science* 260 (1993), 810–812.

[5] A.K. Sood, R. Holmes, M.J. Hendrix and R.E. Buller, Application of the National Cancer Institute international criteria for determination of microsatellite instability in ovarian cancer, *Cancer Res* 61 (2001), 4371–4374.

[6] C.R. Boland, S.N. Thibodeau, S.R. Hamilton, D. Sidransky, J.R. Eshleman, R.W. Burt, S.J. Meltzer, M.A. Rodriguez-Bigas, R. Fodde, G.N. Ranzani and S. Srivastava, A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer, *Cancer Res* 58 (1998), 5248–5257.

[7] I. Tomlinson, S. Halford, L. Aaltonen, N. Hawkins and R. Ward, Does MSI-low exist? *J Pathol* 197 (2002), 6–13.

[8] T.M. Pawlik, C.P. Raut and M.A. Rodriguez-Bigas, Colorectal carcinogenesis: MSI-H versus MSI-L, *Dis Markers* 20 (2004), 199–206.
