Disease Markers 20 (2004) 237–250
IOS Press

Development of a fluorescent multiplex assay for detection of MSI-High tumors

Jeffery W. Bacher[a,*], Laura A. Flanagan[a], Regenia L. Smalley[b], Nadine A. Nassif[a], Lawrence J. Burgart[b], Richard B. Halberg[a], Wael M. Abdel Megid[a] and Stephen N. Thibodeau[b]
[a]*Promega Corporation, Madison, WI, USA*
[b]*Mayo Clinic, College of Medicine, Rochester, MN, USA*

Abstract. Determining whether a tumor exhibits microsatellite instability (MSI) is useful in identifying patients with hereditary non-polyposis colorectal cancer and sporadic gastrointestinal cancers with defective DNA mismatch repair (MMR). The assessment of MSI status aids in establishing a clinical prognosis and may be predictive of tumor response to chemotherapy. A reference panel of five markers was suggested for MSI analysis by a National Cancer Institute (NCI) workshop in 1997 that has helped to standardize testing. But this panel of markers has limitations resulting from the inclusion of dinucleotide markers, which are less sensitive and specific for detection of tumors with MMR deficiencies compared to other types of markers that are currently available. This study demonstrates that mononucleotides are the most sensitive and specific markers for detection of tumors with defects in MMR and identifies an optimal panel of markers for detection of MSI-H tumors. A set of 266 mono-, di-, tetra- and penta-nucleotide repeat microsatellite markers were used to screen for MSI in colorectal tumors. The best markers for detection of MSI-H tumors were selected for a *MSI Multiplex System*, which included five mononucleotide markers: *BAT-25*, *BAT-26, NR-21, NR-24* and *MONO-27*. In addition, two pentanucleotide markers were added to identify sample mix-ups and/or contamination. We classified 153 colorectal tumors using the new MSI Multiplex System and compared the results to those obtained with a panel of 10 microsatellite markers combined with immunohistochemical (IHC) analysis. We observed 99% concordance between the two methods with nearly 100% accuracy in detection of MSI-H tumors. Approximately 5% of the MSI-H tumors had normal levels of four MMR proteins and as a result would have been misclassified based solely on IHC analysis, emphasizing the importance of performing MSI testing. The new *MSI Multiplex System* offers several distinct advantages over other methods of MSI testing in that it is both extremely sensitive and specific and amenable to high-throughput analysis. The *MSI Multiplex System* meets the new recommendations proposed at the recent 2002 NCI workshop on HNPCC and MSI testing and overcomes problems inherent to the original five-marker panel. The use of a single multiplex fluorescent MSI assay reduces the time and costs involved in MSI testing with increased reliability and accuracy and thus should facilitate widespread screening for microsatellite instability in tumors of patients with gastrointestinal cancers.

Keywords: Microsatellite instability, MSI, hereditary non-polyposis colorectal cancer, HNPCC and mononucleotide

1. Introduction

The majority of tumors from patients with hereditary non-polyposis colorectal cancer (HNPCC also known as Lynch syndrome) and approximately 15% of sporadic colorectal tumors exhibit microsatellite instability (MSI), which is characterized by alterations in the length of short tandem repeats in tumor DNA [1,16,44]. Microsatellite instability is caused by DNA mismatch repair (MMR) deficiency and subsequent failure to repair errors introduced during replication of DNA [6,11, 22]. Repetitive DNA sequences like microsatellites are especially prone to errors during DNA replication. The analysis of their stability combined with the analysis of DNA MMR proteins' levels by immunohistochemistry (IHC) provides a sensitive and specific method for identifying tumors with defective DNA mismatch re-

*Corresponding author: Jeff Bacher, Ph D , Promega Corporation, 2800 Woods Hollow Road, Madison, WI, 53711-5399, USA Tel : +1 608 277 2608; E-mail: jeff bacher@promega com

ISSN 0278-0240/04/$17 00 © 2004 – IOS Press and the authors All rights reserved

JHU 2022
Merck Sharp v. Johns Hopkins
IPR2024-00649

pair. The MSI status of a tumor is predictive of clinical outcome and may also be predictive of tumor response to adjuvant therapies [9,13,28–30,35,37,39,47,49].

In 1997, a National Cancer Institute sponsored workshop recommended a panel of five microsatellite markers for detection of MSI consisting of two mononucleotide markers (*BAT-25* and *BAT-26*) and three dinucleotide repeats (*D2S123, D5S346* and *D17S250*). This panel of markers is referred to as the Bethesda panel [5]. Tumors with instability in two or more of these markers are defined as MSI-High (MSI-H), whereas those with one unstable marker are designated as MSI-Low (MSI-L). Tumors with no detectible alterations are stable (MSS). The Bethesda panel of markers provides a uniform set of markers and criteria for tumor classification that has helped standardize MSI analysis. This panel of markers does, however, have some limitations that were detailed at a 2002 NCI workshop that revised recommendations for MSI testing (46). These limitations are a result of including dinucleotide repeats in the panel, as they are less sensitive and less specific for detection of tumors with MMR deficiencies than mononucleotides and can result in misclassification of MSI-L tumors as MSI-H. In addition, interpretation of allelic profiles using dinucleotide repeats is complicated by the presence of stutter, which is a PCR artifact produced by repeat slippage during amplification.

Current evidence suggests that only the MSI-H phenotype is associated with HNPCC and sporadic tumors with MMR defects and that there may be molecular and clinopathological differences between MSI-H and MSS/MSI-L colorectal cancers [references]. As compared to MSS tumors, MSI-H tumors are more likely to occur in individuals with a positive family history of colorectal cancer, to arise in the right colon, to occur in women, and to be associated with a favorable prognosis [25,26,28,32,39,40,44,49]. MSI-H tumors are more likely to have mutations in genes with short repetitive tracts including*TGF-βRII, BAX, IGF2R* genes [27,30, 32,34,42,50] and less likely to have loss of *APC* [1,15, 31,38] or have mutations in p53 [1,7,24,31,38,41] and K-*ras* genes [1,24,31,38]. Tumors with MSI-H phenotype are usually diploid or nearly diploid as opposed to MSS/MSI-L tumors that are typically aneuploid [18]. At the morphological level, MSI-H tumors are characteristically mucinous, poorly differentiated and show lymphocytic infiltration [17,18,20,36].

While the distinction between MSI-H and MSI-L tumors is generally accepted, the separation between MSI-L and MSS groups remains controversial [14, 21,23,45]. Most non-MSI-H colorectal tumors exhibit some low level of instability if enough markers are used. Halford and colleagues reported that 68% of sporadic non-MSI-H tumors displayed instability in at least one marker when 44 microsatellite markers were tested [14]. Similarly, Laiho and colleagues found 79% of *BAT-26* stable tumors were MSI-L using 377 dinucleotide repeats [21]. Further, MSI analysis with different sub-sets of microsatellite markers resulted in identification of almost totally different MSI-L groups. While it was suggested that MSI-L is a real phenomenon and a quantitative trait, no evidence was found of a discrete MSI-L group or association between MSI-L and any clinopathological feature. It remains unclear whether MSI-L is caused by some specific underlying defect or results from the process of carcinogenesis in general, such as random microsatellite slippage accompanied by clonal expansion.

Therefore, we focused our efforts on distinguishing between tumors with the MSI-H phenotype, a well-established subgroup of colorectal cancers with a characterized underlying cause, and non-MSI-H tumors. Our goals were to improve on the existing panel of MSI markers described by Boland and colleagues to significantly increase the throughput and accuracy of MSI tests [5]. We identified an optimal set of markers that provided maximal sensitivity and specificity for MSI-H tumors lacking expression of one or more MMR genes and incorporated them into a multiplex fluorescent assay for simple, rapid and accurate detection of MSI-H tumor phenotype.

2. Material and methods

2.1. DNA samples

The three different sets of colorectal tumor and matching normal samples used in this study were derived from samples collected at the Mayo Clinic (Rochester, MN), with the exception of sample set A that also contained blood and tissue samples obtained through the Cooperative Human Tissue Network (Columbus, OH). Sample set A consisted of 35 MSI-H and 30 non-MSI-H pairs, sample set B consisted of 49 MSI-H and 114 non-MSI-H pairs and sample set C consisted of 72 MSI-H and 81 non MSI-H pairs. DNA from tumor samples was isolated from paraffin-embedded material and corresponding normal DNA was derived from peripheral blood as described elsewhere [23].

Genomic DNA from 177 Caucasian-American, 236 African-American and 127 Asian-American healthy in-

dividuals was obtained from the BODE Technologies Group (Springfield, VA) and genotyped to determine allelic size ranges and allele frequencies for *BAT-25, BAT-26, NR-21, NR-24* and *MONO-27* markers. The frequency of germline mutations in a selected group of microsatellite markers was estimated by analyzing 200–300 healthy parent-child kindreds also obtained from BODE Technologies Group.

2.2. MSI and IHC analysis

The MSI status of all tumor DNA samples analyzed in this study had been previously determined with a panel of 10 microsatellite markers including mononucleotides *BAT-25, BAT-26, BAT-40, BAT34c4*, dinucleotides *D5S346, D17S250, ACTC, D18S55, D10S197* and a complex penta-tetra-mononucleotide repeat marker *MYCL1*. In addition, the status of mismatch repair proteins was determined by immunohistochemistry assays either for MLH1 and MSH2 in sample set A, or for MLH1, MSH2 and MSH6 in samples sets B and C. IHC assay for PMS2 was performed on sample set C in cases where loss of the other MMR proteins was not observed. Immunohistochemical analysis revealed a lack of MLH1, MSH2, MSH6 or PMS2 proteins in 63%, 25%, 9% and 2% of MSI-H tumor samples tested, respectively. MSI assays and IHC analysis were carried out as described by Lindor *et al*. [23].

We evaluated a total of 266 mono-, di-, tetra- and penta-nucleotide repeat microsatellite loci for sensitivity to MSI. Microsatellite markers screened for MSI included: (a) 7 mononucleotide repeats including, *BAT-25, BAT-26, BAT-34c4, BAT-40, NR-21, NR-24* and *MONO-27* [8,43], (b) 6 dinucleotide markers from the Bethesda panel of markers including, *D2S123, D5S346, D17S250, ACTC, D18S55* and *D10S197* [5, 23], (c) 245 tetranucleotide repeat markers from the Research Genetics CHLC/Weber Human Screening Set Version 9.0 (Research Genetics, Huntsville, AL) plus *MYCL1*, and (d) 8 pentanucleotide repeats including, *TP53-penta* and *Penta A, B, C, D, E, F*, and *G* markers [2,3,8]. All markers have been described previously except, *MONO-27* (Genbank accession #AC007684), which was identified by screening Genbank DNA sequence database. PCR and gel electrophoresis of markers from the CHLC/Weber Human Screening Set Version 9.0 were amplified following the manufacturer's protocol. All other markers were amplified in 25 μl PCR reactions using 2.5 μl GoldST★R 10X Buffer (Promega, Madison, WI), 0.1-1 μM each primers, 0.05 μl AmpliTaq Gold DNA Polymerase (5Units/μl) (Perkin Elmer, Wellesley, MA) per locus and 1-2ng DNA. PCR was performed on a PE 9600 Thermal Cycler using cycling profile: 1 cycle 95°C for 11 minutes; 1 cycle 96°C for 1 minute; 10 cycles 94°C for 30 seconds, ramp 68 seconds to 56°C, hold for 30 seconds, ramp 50 seconds to 70°C, hold for 45 seconds; 20 cycles at 90°C for 30 seconds, ramp 60 seconds to 56°C, hold for 30 seconds, ramp 50 seconds to 70°C, hold for 45 seconds; 60°C for 30 minutes final extension; 4°C hold.

Separation and detection of amplified fragments was performed on an ABI PRISM® 310 or 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) following the manufacturer's protocol. Data was analyzed with GeneScan Analysis and Genotyper Software packages from Applied Biosystems to identify predominate allele size for each locus. Allelic patterns or genotypes for normal and tumor pairs were compared and scored as MSI positive if one or more different alleles were present in the tumor DNA samples that were not found in normal samples from the same individual.

The classification of microsatellite instability was based on guidelines suggested by a National Cancer Institute workshop [5]. Using the 5-locus Bethesda panel of microsatellite markers, tumor samples with 40% MSI were classified as MSI-high (MSI-H), less than 40% as MSI-low (MSI-L), and no alterations were classified as microsatellite-stable (MSS). If more than 5 markers are to be used, MSI-H group was defined as tumors having MSI at 30% or more of the markers tested, whereas the MSI-L tumors exhibit MSI in 1–29% of the markers.

2.3. Multiplex analysis

MSI multiplex analysis was performed using 1–2ng of template DNA in a 25 μl PCR reaction using the following protocol: 17 μl nuclease free water, 2.5 μl GoldST★R 10X Buffer (Promega), 2.5 μl 10X multiplex primer mix, 0.5 μl AmpliTaq Gold DNA Polymerase at 5Units/μl and 2.5 μl DNA at 0.4–0.8ng/μl. PCR was performed on a PE 9600 Thermal Cycler using cycling profile: 1 cycle 95°C for 11 minutes; 1 cycle 96°C for 1 minute; 10 cycles 94°C for 30 seconds, ramp for 68 seconds to 56°C, hold for 30 seconds, ramp 50 seconds to 70°C, hold for 45 seconds; 20 cycles at 90°C for 30 seconds, ramp for 60 seconds to 56°C, hold for 30 seconds, ramp 50 seconds to 70°C, hold for 45 seconds; 60°C for 30 minutes final extension; 4°C hold.

Table 1
Sensitivity and specificity of microsatellite markers for MSI-H tumors

Locus	Repeat type	Genbank accession number		Sensitivity	Specificity	Sample set
BAT-40	Mono	M38180		100% (118/118)	95% (188/198)	B & C
MONO-27	Mono	AC007684		98% (107/109)	100% (177/177)	B & C
NR-21	Mono	XM_033393		99% (64/65)	100% (74/74)	C
BAT-25	Mono	U63834		96% (115/120)	99 5% (187/188)	B & C
BAT-26	Mono	U41210		93% (111/120)	100% (194/194)	B & C
NR-24	Mono	X60152		92% (56/61)	100% (73/73)	C
BAT-34c4	Mono	–		92% (102/111)	99 5% (188/189)	B & C
			Mean	96%	99%	
D2S123	Di	Z16551		91% (31/34)	86% (25/29)	A
D18S55	Di	Z16621		90% (88/98)	97% (182/188)	B & C
D10S197	Di	Z16611		89% (88/99)	96% (171/179)	B & C
D17S250	Di	X54562		88% (93/106)	93% (170/182)	B & C
ACTC	Di	NM_005159		84% (93/111)	95% (184/193)	B & C
D5S346	Di	NM_005669		81% (91/112)	98% (188/192)	B & C
			Mean	87%	94%	
D3S2432	Tetra	G08240		76% (32/42)	98% (80/82)	B
D7S3070	Tetra	G27340		74% (35/47)	95% (105/111)	B
D7S3046	Tetra	G10353		68% (27/40)	98% (92/94)	B
D7S1808	Tetra	G08643		67% (31/46)	98% (112/114)	B
D10S1426	Tetra	G08812		61% (28/46)	95% (105/111)	B
			Mean	69%	97%	
Penta-E	Penta	–		57% (16/28)	100% (29/29)	A
Penta-B	Penta	–		43% (12/28)	96% (27/28)	A
Penta-D	Penta	AC000014		35% (30/87)	100% (104/104)	A & C
Penta-A	Penta	–		31% (9/29)	100% (29/29)	A
Penta-F	Penta	–		21% (6/29)	100% (29/29)	A
Penta-G	Penta	Z83847		21% (6/29)	97% (28/29)	A
Penta-C	Penta	AL138752		14% (12/87)	98% (100/102)	A & C
TP53-penta	Penta	–		14% (4/29)	97% (28/29)	A
			Mean	27%	99%	
MYCL1	Complex			80% (94/117)	85% (164/192)	B & C

Sensitivity is defined as the percent of MSI-H samples with MSI
Specificity is defined as the percent non-MSI-H samples in which that marker was stable
Data Set A = MSI-H ($N = 35$), non MSI-H ($n = 30$)
Data Set B = MSI-H ($N = 49$), non MSI-H ($n = 114$)
Data Set C = MSI-H ($N = 74$), non MSI-H ($n = 81$)

Small pool PCR was also performed with the MSI multiplex using 6pg genomic DNA (roughly 1 genome equivalent), as determined by PicoGreen dsDNA Quantitative Kit (Molecular Probes, Eugene OR) following the manufacturer's protocol. The approximate number of genome equivalents were estimated by amplifying increasing amounts of (0.1–1 μl) of a 10pg/μl DNA dilution in a total of 10 PCR reactions, followed by Poisson analysis of the number of reactions positive and negative for a given marker [51]. Small pool PCR was carried out in 10 μl reactions by proportionally reducing the PCR reaction components described above for the MSI Multiplex and by increasing the number of PCR cycles to 35. Capillary electrophoresis and sample analysis was as describe above except, GeneScan (Applied Biosystems) data was imported into a Genotyper software (Applied Biosystems) macro developed for the MSI Multiplex System to automate genotyping of samples, including assigning locus names, allele sizes and to compare matching normal and tumor sample pairs for shifts in allele sizes.

2.4. Allele frequency determinations

Allele size ranges for all loci included in MSI multiplex were determined by genotyping at least 540 normal DNA samples from individuals from three different racial groups (African-American, Asian-American, and Caucasian-American) using the MSI Multiplex System. Genotypes were determined for mononucleotide repeat loci *BAT-25, BAT-26, MONO-27, NR-21* and *NR-24* to calculate allele frequencies and percent heterozygosity.

Table 2
Germline and somatic mutation frequencies

		Germline mutation frequency			Somatic mutation frequency in non-MSI-H tumors		
Locus	Repeat	Number alleles	Number mutants	Mutation rate	Number alleles	Number mutants	Mutation rate
BAT-25	Mono	560	0	$< 1.8 \times 10^{-3}$	576	2	$3\,4 \times 10^{-3}$
BAT-26	Mono	560	0	$< 1.8 \times 10^{-3}$	606	0	$< 1.7 \times 10^{-3}$
MONO-27	Mono	560	0	$< 1.8 \times 10^{-3}$	420	0	$< 2.4 \times 10^{-3}$
BAT-40	Mono	534	2	3.7×10^{-3}	436	10	$2\,3 \times 10^{-2}$
D2S123	Di	534	0	$< 1.9 \times 10^{-3}$	60	4	$6\,7 \times 10^{-2}$
D5S346	Di	534	0	$< 1.9 \times 10^{-3}$	525	4	$7\,6 \times 10^{-3}$
D17S250	Di	534	0	$< 1.9 \times 10^{-3}$	426	13	$3\,1 \times 10^{-2}$
D7S1808	Tetra	560	0	$< 1.8 \times 10^{-3}$	290	2	$6\,9 \times 10^{-3}$
D7S3046	Tetra	560	0	$< 1.8 \times 10^{-3}$	250	4	$1\,6 \times 10^{-2}$
D7S3070	Tetra	560	1	1.8×10^{-3}	284	7	$2\,5 \times 10^{-2}$
D10S1426	Tetra	560	0	$< 1.8 \times 10^{-3}$	284	9	$3\,2 \times 10^{-2}$
MYCL1	Tetra	534	0	$< 1.9 \times 10^{-3}$	386	36	$9\,3 \times 10^{-2}$

2.5. Mutation studies

Germline mutation frequency was determined for selected mononucleotides (*BAT-25, BAT-26, BAT-40, MONO-27*), dinucleotides (*D2S123, D5S346, D17S250*), tetranucleotides (*D3S2432, D7S1808, D7S3046, D7S3070, D10S1426* and *MYCL1*) by pedigree analysis of up to 280 healthy parent-child pairs. Parentage of all parent-child kindreds was confirmed by analysis with PowerPlex®16 (Promega) which provided paternity/maternity indices exceeding a million to one in favor of parentage for each kindred. Somatic mutations frequencies were estimated by MSI analysis of non-MSI-H high tumor DNA samples with normal mismatch repair protein expression.

3. Results

3.1. Screening microsatellite markers

Several criteria were used to select the best markers for MSI analysis, including sensitivity, specificity, ease of interpretation and percent stutter. Sensitivity of a marker was defined as the percent of MSI-H samples in which that marker was unstable, whereas specificity of a marker was defined as the percent of non-MSI-H samples (MSS and MSI-L) in which that marker was stable. Stutter products are PCR artifacts that are generated by polymerase slippage during amplification of short tandem repeat sequences that differ in size from the actual allele by one or more repeat units.

Markers showing the highest sensitivity and specificity for each of the different repeat types are summarized in Table 1. By far the most sensitive and specific markers were the mononucleotides, showing between 92 and 100% sensitivity and 99.5 to 100% specificity for MSI-H samples. The order of sensitivity observed was mononucleotides, dinucleotides, tetranucleotides then pentanucleotides with an average of 96%, 87%, 69% and 27% of MSI-H tumors exhibiting instability, respectively. The ranking for specificity was different, with mononucleotide markers showing an average of 99% specificity (excluding *BAT-40*), followed by pentanucleotide markers at slightly less than 99%, tetranucleotides at 97% and dinucleotides at 94% specificity. Dinucleotide markers *D2S123* and *D17S250* and tetranucleotide marker *MYCL1* exhibited lowest specificities for MSI-H. This observation was also reported by Whitehall *et al*. (2001) who found that *MYCL1* and *D2S123* were the most commonly mutated microsatellite loci in MSI-L cancers, showing MSI in 67% (28/42) and 24% (10/42) of MSI-L tumors, respectively [48].

Two qualities of microsatellite markers that affect the ease of interpreting allelic profiles are the amount of stutter product generated during PCR amplification and heterozygosity. In general, loci with shorter repeat units have a higher incidence of stutter (Fig. 1). For example, mononucleotide markers *BAT-25* and *BAT-26* had stutter peaks at the N-1 position (1bp or one repeat unit less that actual allele peak) that were on average 79% and 72% as high as the actual allele peak. Dinucleotide markers *D5S346, D17S250* and *D2S123* had stutter peaks at N-2 position that were on average 35%, 49% and 53% as high as the actual allele peaks. In comparison, tetranucleotide repeat loci *D10S1426, D7S1808* and *D7S3046* had N-4 stutter peaks averaging of 6–7%, and pentanucleotide loci *Penta-D, Penta-C and Penta-E* had N-5 stutter peaks all averaging less than 1%. The amount of stutter clearly affects ease of scoring as demonstrated by MSI analysis with high stutter dinucleotide and low stutter tetranucleotides (Fig. 2). But unfortunately the more stable tetra- and

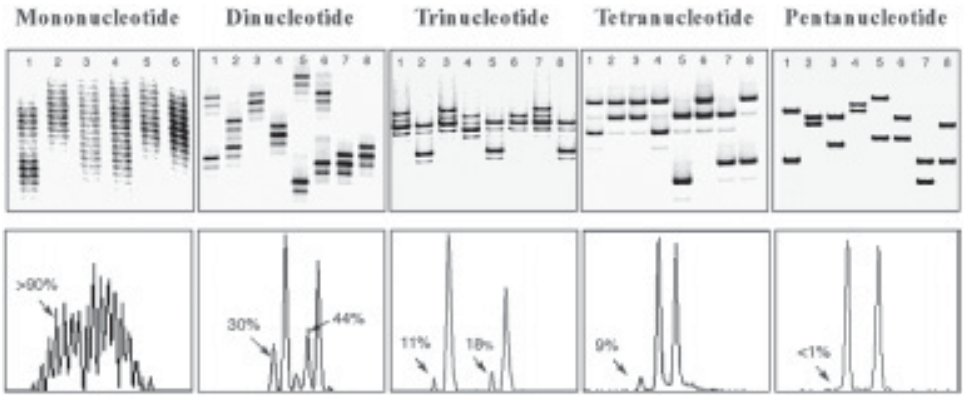

Fig 1 Effect of microsatellite repeat unit length on stutter (repeat slippage) DNA samples were amplified using fluorescently labeled primers for mono- (*BAT-40*), di- (*D5S346*), tri- (*TBP*), tetra- (*D5S818*) and pentanucleotide microsatellite repeat loci (*Penta D*) Stutter is observed more frequently with shorter repeat unit length

pentanucleotide repeats generally displayed lower sensitivity for MSI. Another factor that affects ease of genotyping is heterozygosity. Genotyping DNA from homozygous individuals is easier than from a heterozygous individual. Dinucleotide repeats exhibited both high levels of stutter and heterozygosity. By contrast, mononucleotides were nearly monomorphic and allelic shifts in MSI-H tumors were generally large, making MSI analysis relatively straightforward.

3.2. Somatic and germline mutation rates

Dinucleotide and tetranucleotide repeats were frequently altered in non-MSI-H tumor samples exhibiting normal levels of MMR protein staining in IHC assays. We estimated the frequency of somatic mutations in up to 303 non-MSI-H tumor samples with normal IHC (Table 2). Elevated somatic mutation rates were observed for most di- and tetranucleotide repeats tested and for mononucleotide marker *BAT-40*. The greatest increase in mutation rates were observed in *MYCL1* and *D2S123* markers, with 19% and 13% non-MSI-H tumors showing mutations, respectively.

We speculated the reduced specificity of some microsatellite repeats reflects higher spontaneous mutation rates of these markers. We tested this hypothesis by genotyping 208 parent-child kindreds for germline mutations in mononucleotide markers (*BAT-25, BAT-26, BAT-40* and *MONO-27*), dinucleotide markers (*D2S123, D5S346* and *D17S250*), tetranucleotide markers (*D3S2432, D7S1808, D7S3046, D7S307* and *D10S1426*) and one complex penta/tetra/mono marker (*MYCL1*). One germline mutation was observed out of 560 alleles at tetranucleotide marker *D7S3046* and two mutations out of 534 alleles were observed at mononucleotide marker *BAT-40* (Table 2). This data does not support an association between high somatic mutation frequency in non-MSI-H tumors and higher germline mutations frequencies, except for *BAT-40*. Germline hypermutability of the *BAT-40* poly-A tract was also observed by Bacon and colleagues (2001) who reported mutations in 7% of germline allele transmissions and suggested that stable markers may be too sensitive to provide the specificity needed for MSI testing [4].

We predict that MSI-L tumors could be misclassified as MSI-H if microsatellite markers with lower specificity are used. In our study the average percent of dinucleotide markers found mutated in non-MSI-H tumors with normal IHC was 6% (Table 1). Thus, the estimated background mutation frequency for any two of these dinucleotide markers in all colorectal tumors is approximately 0.4%. At this background mutation frequency, around 3% of MSI-H tumors could be misclassified using a panel of five markers containing two dinucleotide markers with 94% specificity. To test this hypothesis, we analyzed a set of colorectal tumors (sample set A) using the Bethesda panel and observed 1 out of 30 (3%) non-MSI-H tumor samples with mutations in both *D2S123* and *D17S250*. This sample would have been incorrectly misclassified as MSI-H using the Bethesda guidelines as indicated by subsequent analysis showing stability at nine other markers and normal IHC. By contrast, the chance of misclassifying MSI-L samples as MSI-H using the five mononucleotide markers in the MSI Multiplex System based on our specificity data is virtually nonexistent, since only *BAT-25* was altered in non-MSI-H samples and this occurred in only 0.5% (1/188) of the samples.

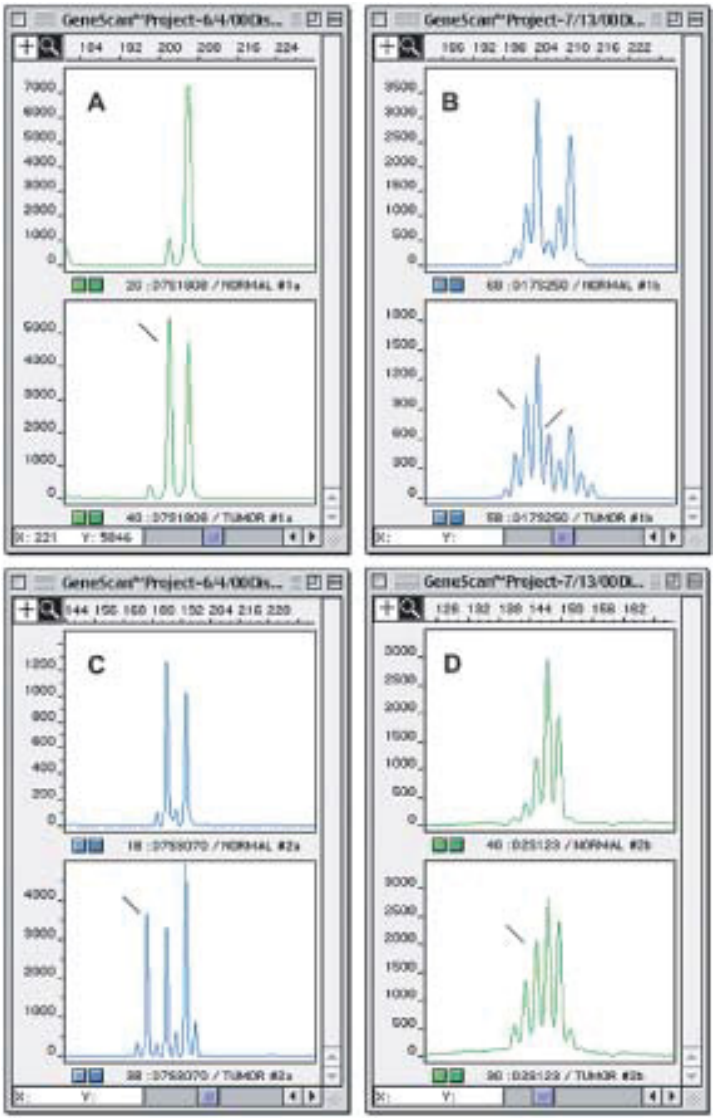

Fig 2 MSI analysis with dinucleotide and tetranucleotide markers Allelic profiles resulting from amplification of tetranucleotide repeat markers were simpler to interpret than were those of dinucleotide markers because of lower stutter Matching normal and MSI-H tumor samples were amplified with tetranucleotide marker *D7S1808* (A) and dinucleotide marker *D17S250* (B) loci A second matching pair of normal and MSI-H tumor samples were amplified with tetranucleotide *D7S3070* (C) and dinucleotide *D2S123* (D) marker New alleles present in MSI-H sample (bottom panels) are indicated with an arrow High stutter in dinucleotide markers complicated MSI analysis, especially when shifted alleles were the same size as stutter peaks

3.3. Construction and evaluation of the fluorescent MSI Multiplex System

Five mononucleotide repeat markers were selected for multiplexing based upon the characteristics of the different classes of repeat markers and the variation in the sensitivity and specificity of particular markers. Information about the mononucleotide markers *BAT-25, BAT-26, NR-21, NR-24* and *MONO-27* included in the *MSI Multiplex System* is given in Table 3. These five mononucleotide markers were found to be the most sensitive and specific for detection of MSI-H tumors. In addition, these markers were found to be nearly monomorphic, which simplifies interpretation

of allelic patterns when only a single allele is present in normal DNA and offers the possibility of performing MSI assays without the need for matching normal samples [43]. Two polymorphic pentanucleotide repeats, *Penta-C* and *Penta-D*, were included in multiplex to detect potential sample mix-ups and/or contamination problems that might not be apparent when using only the nearly monomorphic markers. A matching probability between 1:10 and 1:16 was calculated for *Penta-C*, and between 1:18 and 1:33 for *Penta-D* based upon allele frequency data [2]. Matching probability is the average number of people you would have to survey before you would find the same allelic pattern as a randomly selected individual. Using both pentanucleotide markers increases the matching probability such that only 0.2 to 0.6% of individuals will have the same allelic profile (depending upon ethnic background). Polymorphic pentanucleotide markers were selected instead of di- and tetranucleotides because pentanucleotides are the most stable repeat type in mismatch repair deficient cells, making sample identification less ambiguous. The pentanucleotide repeat markers were not used in determining MSI status.

PCR primers for the multiplex were designed to allow robust co-amplification all markers in a single PCR reaction without generation of non-specific products. Fluorescently labeled PCR products were resolved by size and color by fluorescent capillary electrophoresis with AB 310 or 3100 Genetic Analyzers. An example of a GeneScan (Applied Biosystems) electropherogram showing a typical allelic pattern resulting from amplification of normal DNA or MSS tumors with the MSI Multiplex System is shown in Fig. 3. All PCR products observed for the mononucleotide markers were less than 160bp, allowing robust amplification of fragmented DNA extracted from most paraffin-embedded tissues.

3.4. Allele size ranges and frequencies

Genomic DNA from 177 Caucasian-American, 236 African-American and 127 Asian-American healthy individuals was genotyped to determine allelic size ranges to confirm there was no overlap between markers (Table 3). The common allele size for *NR-21, BAT-26, BAT-25, NR-24* and *MONO-27* markers was 98, 113, 120, 130 and 150bp, respectively. PCR products ranged from 98 to 160bp for all mononucleotide markers, and between 137 and 209bp for all pentanucleotide markers. There was sufficient size difference between markers to prevent overlap even with shortened alleles observed in MSI-H samples.

We examined allele frequencies and heterozygosity levels for MSI multiplex makers in 540 individuals to determine whether they were sufficiently monomorphic to distinguish between mutations and polymorphisms if only tumor samples were used in MSI analysis (Table 4). The most polymorphic population examined was the African-American group and, therefore, this ethnic background would have the highest likelihood of being heterozygous for the mononucleotide markers [33,43]. The percent heterozygosity for *BAT-25, BAT-26, MONO-27, NR-21* and *NR-24* markers in the African-American population ($n = 236$) was 16.1%, 10.9%, 2.5%, 2.2% and 0.9%, respectively. Based on these data, the highest proportion of African-American individuals that are heterozygous for two mononucleotide markers, *BAT-25* and *BAT-26*, would be 1.8%. The highest proportion of individuals heterozygous for the three markers, *BAT-25, BAT-26* and *MONO-27*, would be 0.04%. When we only considered heterozygotes that differed by $\leqslant$ 3bp, the percent heterozygosity values lowered to 11.5%, 9.6%, 0.9%, 1.3% and 0.9% for *BAT-25, BAT-26, MONO-27, NR-21* and *NR-24*, respectively. This reduced the predicted frequency of individuals heterozygous for two markers, *BAT-25* and *BAT-26*, to 1.1%, and for three markers, *BAT-25, BAT-26* and *NR-24*, to 0.014%.

The majority of allelic variation in mononucleotide markers was within 3bp of the most common allele. To test whether this variation reflected real differences is allele sizes, or was due to run-to-run differences in sizing, we looked at consistency of sizing between multiple ABI 3100 runs of the same PCR product. Allele sizes were uniform across runs with the exception of a small number of alleles that had two major peaks of similar height differing in size by 1bp. Small pool PCR analysis of the ambiguously sized alleles revealed that small variations in sizing of mononucleotide markers can occur in individuals heterozygous for alleles that differ in size by a few base pairs and we hypothesize that random events could result in unequal amplification and subsequent variation in the prominent allele. Thus, only shifts of 3 bp or more should be considered mutants to avoid confusion with natural low-level polymorphic variation in this type of marker.

3.5. Determination of MSI tumor phenotype with the MSI Multiplex System

Multiplex MSI analysis was performed on a set of 72 MSI-H and 81 non MSI-H matching tumor/normal

Table 3
PCR product sizes and respective fluorescent dyes for loci contained in the MSI Multiplex System

Locus name	Genebank number	Gene	Repeat motif	Size range (bp)	Fluorescent label
NR-21	XM_033393	SLC7A8	$(A)_{21}$	94–101	JOE
BAT-26	U41210	MSH2	$(A)_{26}$	103–115	FL
BAT-25	L04143	c-kit	$(A)_{25}$	113–124	JOE
NR-24	X60152	ZNF-2	$(A)_{24}$	130–133	TMR
MONO-27	AC007684	MAP4K3	$(A)_{27}$	142–154	JOE
Penta C	AL138752	–	$(AAAAG)_{3-15}$	143–194	TMR
Penta D	AC000014	–	$(AAAAG)_{2-17}$	135–201	FL

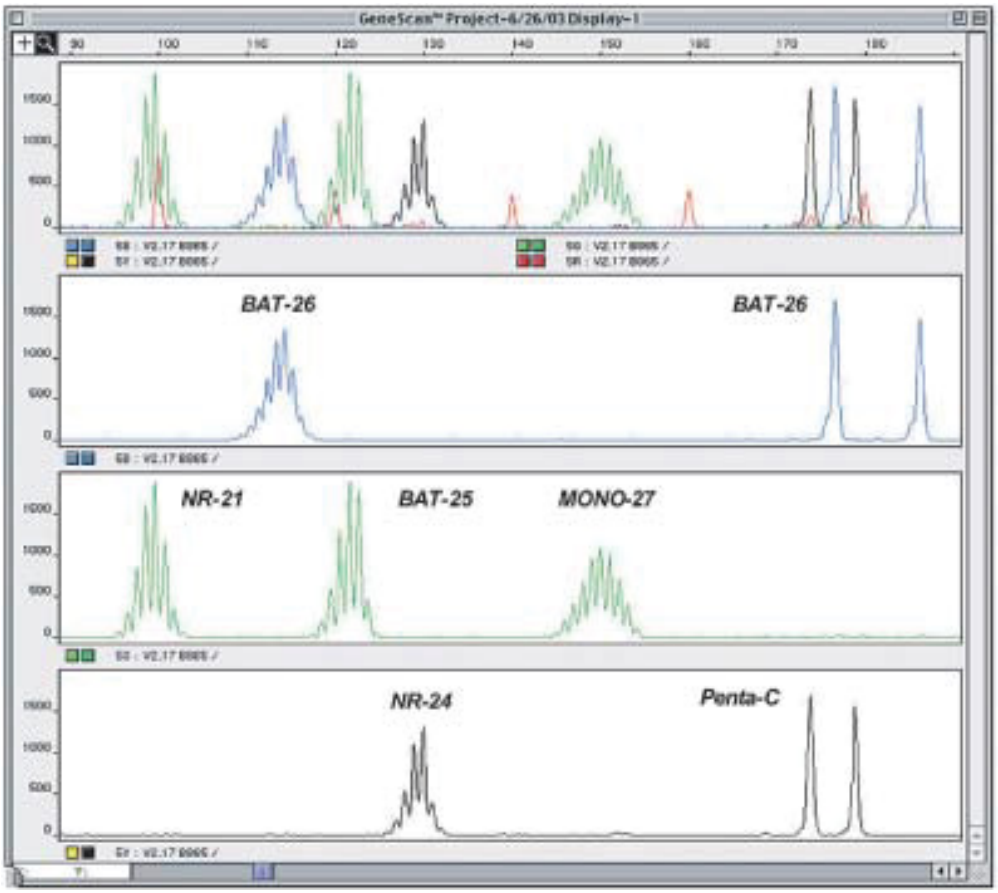

Fig 3 MSI analysis of normal DNA samples with the *MSI Multiplex System* The fluorescent *MSI Multiplex System* contains five nearly monomorphic mononucleotide repeat loci (*BAT-25, BAT-26, MONO-27, NR-21* and *NR-24*) that are very sensitive and specific for instability in MSI-H tumors In addition, two polymorphic pentanucleotide repeats (*Penta-C* and *Penta-D*) are included as internal controls to identify sample mix-ups and/or contamination problems One nanogram of template DNA was amplified using fluorescently labeled primers in a single PCR reaction and products analyzed on an ABI 3100 capillary electrophoresis instrument

pairs (sample set C) that had been previously characterized for MSI status using a panel of 10 microsatellite markers and by IHC analysis for MLH1, MSH2, MSH6 and PMS2 A typical example of shifts in allele sizes observed in MSI-H samples is shown in Fig. 4. There was 97% (70/72) concordance between the MSI classification using the *MSI Multiplex System* and the 10-marker panel for the MSI-H group of specimens. One of these samples was unstable for 4 of 8 informative markers with the 10-marker panel (3 dinucleotide markers and *BAT-40*) and exhibited no instability with the *MSI Multiplex System*. This sample had normal

Table 4
Allele Frequencies for mononucleotide markers in the MSI Multiplex System

NR-21			
Size (bp)	Caucasian	African	Asian
94	1	2	8
95	1	0	0
96	0	3	0
97	0	60	0
98	27	100	14
99	134	62	89
100	14	9	23
101	1	0	1
Total	178	236	135

BAT-26			
Size (bp)	Caucasian	African	Asian
103	0	23	0
104	0	0	0
105	0	0	0
106	0	2	1
107	1	0	0
108	0	0	0
109	0	0	0
110	0	0	0
111	0	4	0
112	14	38	3
113	140	131	115
114	22	57	8
115	0	0	0
Total	177	255	127

NR-24			
Size (bp)	Caucasian	African	Asian
130	60	197	49
131	113	34	77
132	0	0	0
133	0	2	2
Total	173	233	128

MONO-27			
Size (bp)	Caucasian	African	Asian
142	0	1	0
143	0	0	0
144	0	0	0
145	0	0	0
146	0	0	0
147	0	0	0
148	0	3	0
149	59	88	42
150	107	141	83
151	5	4	1
152	1	1	0
153	0	0	0
154	0	1	0
155	0	0	1
Total	172	239	127

BAT-25			
Size (bp)	Caucasian	African	Asian
113	1	0	0
114	0	1	0
115	0	12	0
116	0	7	0
117	0	1	0
118	0	8	0
119	0	14	0
120	7	111	4
121	108	103	93
122	58	12	30
123	8	5	0
124	0	0	0
Total	182	274	127

IHC for all four MMR proteins analyzed. The second discordant sample demonstrated instability in 3 of 10 informative markers with the 10-marker panel (two dinucleotide markers and *BAT-40*) but did not show instability at any of the markers in the MSI Multiplex System panel. Repeat analysis of the MSI testing confirmed the original results. Additionally, this sample was negative for MLH1, but the IHC could not be repeated because tissue sections were no longer available. A sub-group of the MSI-H samples from sample set C showed normal IHC for all MMR proteins tested, but exhibited high levels of MSI with the MSI Multiplex System (100% of markers were altered in all samples) and the 10-marker panel (over 50% markers altered). This data indicates that approximately 5% (4/72) of these samples would have been misclassified if only IHC analysis had been performed. All samples classified as MSI-L with the 10-marker panel were stable when analyzed with the *MSI Multiplex System* and all exhibited normal MMR protein staining with IHC. In fact, no samples were classified as MSI-L using the *MSI Multiplex System*. All of the MSS tumor/normal pairs were correctly classified. Thus, the *MSI Multiplex System* accurately distinguishes between MSI-H and MSS/MSI-L tumors.

4. Conclusion

This study confirms the observation that mononucleotide repeats are most sensitive and specific markers for MSI analysis [8]. Overall, the order of sensitivity we observed was mononucleotides, dinucleotides, tetranucleotides then pentanucleotides with an average of 96%, 87%, 69% and 27% of MSI-H samples showing instability, respectively. The ranking for specificity was mononucleotide (99%), pentanucleotide (99%), tetranucleotides (97%) and then dinucleotides (94%).

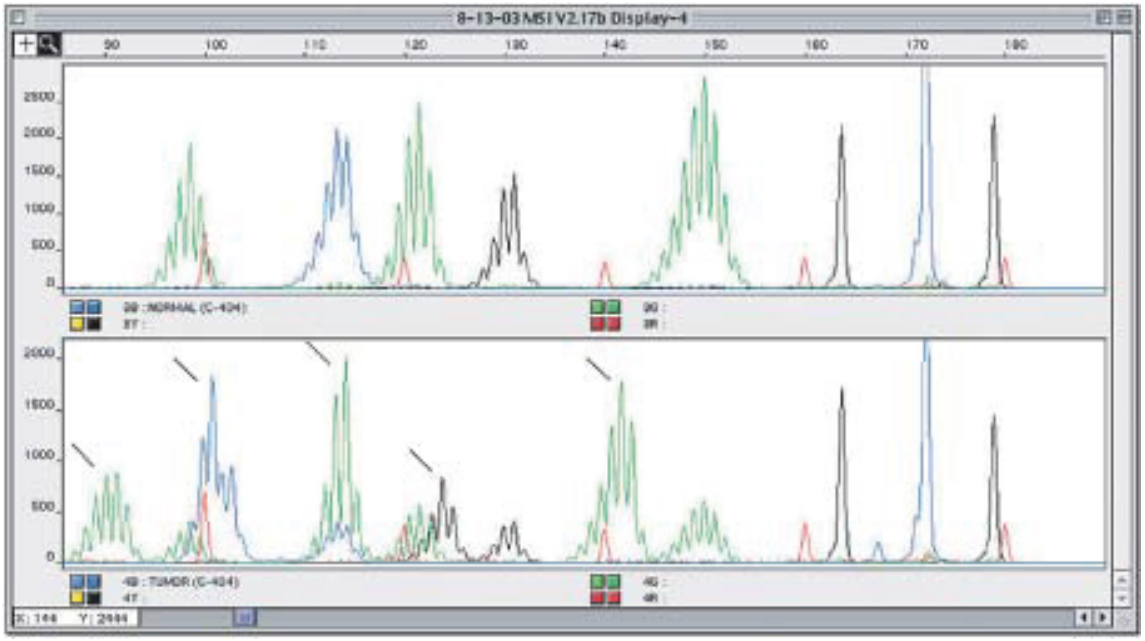

Fig 4 MSI analysis of colorectal tumor and corresponding normal DNA with the *MSI Multiplex System* Electropherogram showing an allelic profile generated from a normal sample (top panel) or from a matching MLH1-deficient tumor sample (bottom panel) using the *MSI Multiplex System* New alleles present in tumor sample that are not found in matching normal sample indicate MSI (new allele peaks indicated by arrows) Notice that 100% of the mononucleotide repeats were unstable while the pentanucleotide repeats remained unchanged The pentanucleotide repeats are not used for MSI classifications, only for the detection of sample mix-ups or contamination problems

Markers exhibiting the lowest specificity for MSI-H tumors were mononucleotide *BAT-40*, dinucleotide markers *D2S123* and *D17S250* and tetranucleotide marker *MYCL1*. Our data also supports the conclusion by Dietmaier et al. (1997) that dinucleotides are most difficult, and that mononucleotides are the easiest type of marker with which to interpret allelic changes in MSI-H tumors [8].

The five most sensitive and specific markers for detection of MSI-H tumors were selected from a panel of 266 mono-, di-, tetra- and penta-nucleotide repeat microsatellite markers. They included mononucleotide markers *BAT-25, BAT-26, NR-21, NR-24* and *MONO-27*, which were combined into a single fluorescent multiplex assay for rapid and accurate determination of tumor MSI tumor phenotype. The sensitivity for the individual markers ranged from 92% to 98% and specificity from 99.5% to 100% (Table 1). We classified 153 colorectal tumors using the new MSI Multiplex System and compared the results to those obtained with a panel of 10 microsatellite markers and immunohistochemical (IHC) analysis. Over 97% (70/72) of MSI-H samples were correctly identified and overall there was 99% (151/153) concordance in MSI classification between the two methods. The reasons for the possible discordant cases are unclear at this time. Approximately 5% of the tumors that were classified as MSI-H using the new MSI Multiplex System exhibited no loss of any MMR proteins by IHC. This value is in close agreement with larger studies conducted at the Mayo Clinic Rochester, where 4 to 5% of MSI-H samples were found to exhibit normal IHC results for MLH1, MSH2, MSH6 and PMS2 (S. Thibodeau, unpublished data). These observations indicate MSI testing may identify MSI-H tumors that are misclassified based only on IHC. The most striking difference between the two marker panels was in classification of the MSI-L group. All of the 43 samples classified as MSI-L using the 10-marker panel (unstable at dinucleotide or *BAT-40* markers only) were scored as MSS with the *MSI Multiplex System*, indicating the high specificity of the MSI multiplex markers for instability in MSI-H tumors only.

Use of mononucleotide markers for MSI testing overcomes one of the disadvantages of the Bethesda panel, that is, the misclassification of MSI-L as MSI-H in cases where only dinucleotide markers are altered. This misclassification can occur because instability of some

of the dinucleotide markers included in the Bethesda panel is not specific to tumors with MMR deficiencies. We found that around 6% of non-MSI-H tumors with normal IHC carried a mutation in a single dinucleotide repeat. Based on our data we estimate that as much as 3% of MSI-H tumors could be misclassified using the Bethesda panel. In fact, we did find that 1 out of 30 non-MSI-high tumors (sample set A) exhibited mutations in dinucleotide markers *D2S123* and *D17S250* from the Bethesda panel. This potential problem was recognized at the 2002 NCI workshop and it was suggested that mononucleotide markers could be used instead of dinucleotide markers to increase sensitivity and specificity [46].

Suraweera and colleagues have proposed that MSI status of a tumor can be assessed without reference to matching normal DNA if a panel of five nearly monomorphic mononucleotide markers is used [43]. Any deviation of 3 or more base pairs in the tumor allele size from the commonly occurring alleles found in the larger population was considered MSI. MSI in 0 to 2 markers was scored as MSS and in 3 or more markers as MSI-H. In general, our allele frequency data for African-American populations supports the conclusion that instability in 3 out of 5 markers in MSS tumors would only occur at very low frequencies (=0.014% false positives) using the makers in our *MSI Multiplex System*. But, the problem of dismissing an actual MSI-H case by defining MSS as less or equal to 2 out of 5 markers unstable was observed. In 74 MSI-H cases, we found one MSI-H case (1.4% false negatives) where only 2 out of 5 mononucleotide makers were unstable. False negatives could be avoided, however, in the small number of cases with only two unstable markers by testing matching normal sample to determine if changes in allele sizes were due to polymorphism rather than mutations. The use of a panel of five nearly monomorphic mononucleotide markers for MSI testing of tumor samples without comparison to matching normal samples has been included into the updated NCI workshop recommendations for MSI testing [46].

Determination of MSI status has become a useful tool for screening for HNPCC and sporadic MSI-H tumors, and defines a subset of colorectal cancers with distinctive molecular and clinopathological features. This study provides strong evidence that mononucleotides are the most sensitive and specific markers for detection of MSI-H and identifies an optimal panel of markers that detects MSI-H tumors with nearly 100% accuracy. The *MSI Multiplex System* was designed to analyze small amounts of DNA isolated from paraffin-embedded tumor samples alone or as matching tumor/normal pairs. The selected panel of markers includes five nearly monomorphic mononucleotides for MSI determinations and two polymorphic pentanucleotide for detection of sample mix-ups and/or contamination problems. This configuration of markers in a single multiplex assay offers several distinct advantages over other methods of MSI analysis in that it provides a method that offers both extremely high sensitivity and specificity while still being amenable to automated high-throughput analysis using existing fluorescent capillary electrophoresis technology. The *MSI Multiplex System* meets the new recommendations proposed at the recent 2002 NCI workshop on HNPCC and MSI testing and overcomes problems inherent to the original five-marker panel. This approach should facilitate widespread screening for microsatellite instability in tumors of patients with gastrointestinal cancers.

Acknowledgements

Supported by grant number CA-76834 from the National Institutes of Health. Tissue samples were provided by the Cooperative Human Tissue Network.

References

[1] 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), 812–816

[2] J Bacher, L F Hennes, T Gu, A Tereba, K A Micka, C J Sprecher, A M Lins, E A Amiott, D R Rabbach, J A Taylor, C Helms, H Donis-Keller and J W Schumm, Pentanucleotide Repeats: Highly Polymorphic Genetic Markers Displaying Minimal Stutter Artifact, Proceedings from the Ninth International Symposium on Human Identification (1999) from http://www promega com/geneticidproc/ussymp9proc/default htlm

[3] J Bacher and J Schumm, Development of Highly Polymorphic Pentanucleotide Tandem Repeat Loci with Low Stutter, *Profiles in DNA* **2** (1998), 3–6

[4] A L Bacon, M G Dunlop and S M Farrington, Hypermutability at a poly(A/T) tract in the human germline, *Nucleic Acids Res* **29** (2001), 4405–4413

[5] 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

[6] C E Bronner, S M Baker, P T Morrison, G Warren, L G Smith, M K Lescoe, M Kane, C Earabino, J Lipford and A Lindblom, Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer, *Nature* **368** (1994), 258–261

[7] P H Cottu, F Muzeau, A Estreicher, J F Flejou, R Iggo, G Thomas and R Hamelin, Inverse correlation between RER+ status and p53 mutation in colorectal cancer cell lines, *Oncogene* **13** (1996), 2727–2730

[8] W Dietmaier, S Wallinger, T Bocker, F Kullmann, R Fishel and J Ruschoff, Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression, *Cancer Res* **57** (1997), 4749–4756

[9] T B Edmonston, K H Cuesta, S Burkholder, A Barusevicius, D Rose, A J Kovatich, B Boman, R Fry, R Fishel and J P Palazzo, Colorectal carcinomas with high microsatellite instability: defining a distinct immunologic and molecular entity with respect to prognostic markers, *Hum Pathol* **31** (2000), 1506–1514

[10] H Elsaleh, D Joseph, F Grieu, N Zeps, N Spry and B Iacopetta, Association of tumour site and sex with survival benefit from adjuvant chemotherapy in colorectal cancer, *Lancet* **355** (2000), 1745–1750

[11] R Fishel, M K Lescoe, M R Rao, N G Copeland, N A Jenkins, J Garber, M Kane and R Kolodner, The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer, *Cell* **77** (1994), 167

[12] J Gayet, X P Zhou, A Duval, S Rolland, J M Hoang, P Cottu and R Hamelin, Extensive characterization of genetic alterations in a series of human colorectal cancer cell lines, *Oncogene* **20** (2001), 5025–5032

[13] R Gryfe, H Kim, E T Hsieh, M D Aronson, E J Holowaty, S B Bull, M Redston and S Gallinger, Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer, *N Engl J Med* **342** (2000), 69–77

[14] S Halford, P Sasieni, A Rowan, H Wasan, W Bodmer, I Talbot, N Hawkins, R Ward and I Tomlinson, Low-level microsatellite instability occurs in most colorectal cancers and is a nonrandomly distributed quantitative trait, *Cancer Res* **62** (2002), 53–57

[15] J Huang, N Papadopoulos, A J McKinley, S M Farrington, L J Curtis, A H Wyllie, S Zheng, J K Willson, S D Markowitz, P Morin, K W Kinzler, B Vogelstein and M G Dunlop, APC mutations in colorectal tumors with mismatch repair deficiency, *Proc Natl Acad Sci USA* **93** (1996), 9049–9054

[16] 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

[17] J R Jass, K A Do, L A Simms, H Iino, C Wynter, S P Pillay, J Searle, G Radford-Smith, J Young and B Leggett, Morphology of sporadic colorectal cancer with DNA replication errors, *Gut* **42** (1998), 673–679

[18] J R Jass, V L Whitehall, J Young and B A Leggett, Emerging concepts in colorectal neoplasia, *Gastroenterology* **123** (2002), 862–876

[19] J R Jass, V Whitehall, J Young and B Leggett, Correspondence re: P Laiho et al , Low-level microsatellite instability in most colorectal carcinomas, *Cancer Res* **62** (2002), 1166–1170 *Cancer Res* **62** (2002), 5989–5990; author reply 5989–5990

[20] H Kim, J Jen, B Vogelstein and S R Hamilton, Clinical and pathological characteristics of sporadic colorectal carcinomas with DNA replication errors in microsatellite sequences, *Am J Pathol* **145** (1994), 148–156

[21] P Laiho, V Launonen, P Lahermo, M Esteller, M Guo, J G Herman, J P Mecklin, H Jarvinen, P Sistonen, K M Kim, D Shibata, R S Houlston and L A Aaltonen, Low-level microsatellite instability in most colorectal carcinomas, *Cancer Res* **62** (2002), 1166–1170

[22] F S Leach, N C Nicolaides, N Papadopoulos, B Liu, J Jen, R Parsons, P Peltomaki, P Sistonen, L A Aaltonen and M Nystrom-Lahti, Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer, *Cell* **75** (1993), 1215–1225

[23] N M Lindor, L J Burgart, O Leontovich, R M Goldberg, J M Cunningham, D J Sargent, C Walsh-Vockley, G M Petersen, M D Walsh, B A Leggett, J P Young, M A Barker, J R Jass, J Hopper, S Gallinger, B Bapat, M Redston and S N Thibodeau, Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors, *J Clin Oncol* **20** (2002), 1043–1048

[24] L Losi, M Ponz de Leon, J Jiricny, C Di Gregorio, P Benatti, A Percesepe, R Fante, L Roncucci, M Pedroni and J Benhattar, K-ras and p53 mutations in hereditary non-polyposis colorectal cancers, *Int J Cancer* **74** (1997), 94–96

[25] R A Lothe, P Peltomaki, G I Meling, L A Aaltonen, M Nystrom-Lahti, L Pylkkanen, K Heimdal, T I Andersen, P Moller and T O Rognum, Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history, *Cancer Res* **53** (1993), 5849–5852

[26] S R Malkhosyan, H Yamamoto, Z Piao and M Perucho, Late onset and high incidence of colon cancer of the mutator phenotype with hypermethylated hMLH1 gene in women, *Gastroenterology* **119** (2000), 598

[27] S Markowitz, J Wang, L Myeroff, R Parsons, L Sun, J Lutterbaugh, R S Fan, E Zborowska, K W Kinzler and B Vogelstein, Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability, *Science* **268** (1995), 1336–1338

[28] L Messerini, M Ciantelli, S Baglioni, A Palomba, G Zampi and L Papi, Prognostic significance of microsatellite instability in sporadic mucinous colorectal cancers, *Hum Pathol* **30** (1999), 629–634

[29] K Mure and T G Rossman, Reduction of spontaneous mutagenesis in mismatch repair-deficient and proficient cells by dietary antioxidants, *Mutat Res* **480–481** (2001), 85–95

[30] C Oliveira, R Seruca, M Seixas and M Sobrinho-Simoes, The clinicopathological features of gastric carcinomas with microsatellite instability may be mediated by mutations of different target genes: a study of the TGFbeta RII, IGFII R, and BAX genes, *Am J Pathol* **153** (1998), 1211–1219

[31] S Olschwang, R Hamelin, P Laurent-Puig, B Thuille, Y De Rycke, Y J Li, F Muzeau, J Girodet, R J Salmon and G Thomas, Alternative genetic pathways in colorectal carcinogenesis, *Proc Natl Acad Sci USA* **94** (1997), 12122–12127

[32] M Planck, E Wenngren, A Borg, H Olsson and M Nilbert, Somatic frameshift alterations in mononucleotide repeat-containing genes in different tumor types from an HNPCC family with germline MSH2 mutation, *Genes Chromosomes Cancer* **29** (2000), 33–39

[33] R Pyatt, R B Chadwick, C K Johnson, C Adebamowo, A de la Chapelle and T W Prior, Polymorphic variation at the BAT-25 and BAT-26 loci in individuals of African origin Implications for microsatellite instability testing, *Am J Pathol* **155** (1999), 349–353

[34] N Rampino, H Yamamoto, Y Ionov, Y Li, H Sawai, J C Reed and M Perucho, Somatic frameshift mutations in the

BAX gene in colon cancers of the microsatellite mutator phenotype, *Science* **275** (1997), 967–969

[35] C M Ribic, D J Sargent, M J Moore, S N Thibodeau, A J French, R M Goldberg, S R Hamilton, P Laurent-Puig, R Gryfe, L E Shepherd, D Tu, M Redston and S Gallinger, Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer, *N Engl J Med* **349** (2003), 247–257

[36] M Risio, G Reato, P F di Celle, M Fizzotti, F P Rossini and R Foa, Microsatellite instability is associated with the histological features of the tumor in nonfamilial colorectal cancer, *Cancer Res* **56** (1996), 5470–5474

[37] J Ruschoff, S Wallinger, W Dietmaier, T Bocker, G Brockhoff, F Hofstadter and R Fishel, Aspirin suppresses the mutator phenotype associated with hereditary nonpolyposis colorectal cancer by genetic selection, *Proc Natl Acad Sci USA* **95** (1998), 11301–11306

[38] S Salahshor, U Kressner, L Pahlman, B Glimelius, G Lindmark and A Lindblom, Colorectal cancer with and without microsatellite instability involves different genes, *Genes Chromosomes Cancer* **26** (1999), 247–252

[39] W S Samowitz, K Curtin, K N Ma, D Schaffer, L W Coleman, M Leppert and M L Slattery, Microsatellite instability in sporadic colon cancer is associated with an improved prognosis at the population level, *Cancer Epidemiol Biomarkers Prev* **10** (2001), 917–923

[40] S Senba, F Konishi, T Okamoto, H Kashiwagi, K Kanazawa, M Miyaki, M Konishi and T Tsukamoto, Clinicopathologic and genetic features of nonfamilial colorectal carcinomas with DNA replication errors, *Cancer* **82** (1998), 279–285

[41] L A Simms, G Radford-Smith, K G Biden, R Buttenshaw, M Cummings, J R Jass, J Young, S J Meltzer and B A Leggett, Reciprocal relationship between the tumor suppressors p53 and BAX in primary colorectal cancers, *Oncogene* **17** (1998), 2003–2008

[42] R F Souza, R Appel, J Yin, S Wang, K N Smolinski, J M Abraham, T T Zou, Y Q Shi, J Lei, J Cottrell, K Cymes, K Biden, L Simms, B Leggett, P M Lynch, M Frazier, S M Powell, N Harpaz, H Sugimura, J Young and S J Meltzer, Microsatellite instability in the insulin-like growth factor II receptor gene in gastrointestinal tumours, *Nat Genet* **14** (1996), 255–257

[43] N Suraweera, A Duval, M Reperant, C Vaury, D Furlan, K Leroy, R Seruca, B Iacopetta and R Hamelin, Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR, *Gastroenterology* **123** (2002), 1804–1811

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

[45] I P Tomlinson, R Hampson, P Karran and W F Bodmer, DNA mismatch repair in lymphoblastoid cells from hereditary nonpolyposis colorectal cancer (HNPCC) patients is normal under conditions of rapid cell division and increased mutational load, *Mutat Res* **383** (1997), 177–182

[46] A Umar, C R Boland, J P Terdiman, S Syngal, A de la Chapelle, J Rüschoff, R Fishel, N M Lindor, L J Burgart, R Hamelin, S R Hamilton, J Jass, A Lindblom, H T Lynch, P Peltomaki, S D Ramsey, M A Rodriguez-Bigas, H F Vasen, E T Hawk, J C Barrett, A N Freedman and S Srivastava, Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability, *J. Natl Cancer Inst* **96** (2004), 261–268

[47] T Watanabe, T T Wu, P J Catalano, T Ueki, R Satriano, D G Haller, A B Benson, 3rd and S R Hamilton, Molecular predictors of survival after adjuvant chemotherapy for colon cancer, *N Engl J Med* **344** (2001), 1196–1206

[48] V L Whitehall, C V Wynter, M D Walsh, L A Simms, D Purdie, N Pandeya, J Young, S J Meltzer, B A Leggett and J R Jass, Morphological and molecular heterogeneity within nonmicrosatellite instability-high colorectal cancer, *Cancer Res* **62** (2002), 6011–6014

[49] C M Wright, O F Dent, M Barker, R C Newland, P H Chapuis, E L Bokey, J P Young, B A Leggett, J R Jass and G A Macdonald, Prognostic significance of extensive microsatellite instability in sporadic clinicopathological stage C colorectal cancer, *Br J Surg* **87** (2000), 1197–1202

[50] H Yamamoto, H Sawai, T K Weber, M A Rodriguez-Bigas and M Perucho, Somatic frameshift mutations in DNA mismatch repair and proapoptosis genes in hereditary nonpolyposis colorectal cancer, *Cancer Res* **58** (1998), 997–1003

[51] C L Yauk, Y E Dubrova, G R Grant and A J Jeffreys, A novel single molecule analysis of spontaneous and radiation-induced mutation at a mouse tandem repeat locus, *Mutat Res* **500** (2002), 147–156