Microsatellite Instability in Saliva from Patients with Hereditary Non-polyposis Colon Cancer and Siblings Carrying Germline Mismatch Repair Gene Mutations

Peter Hu1,2, Chang Woo Lee1, Jing P. Xu2, Crystal Simien1,2, Chuan Li Fan1, Michael Tam1, Louis Ramagli2, Barry W. Brown3, Patrick Lynch4, Marsha L. Frazier5, Michael J. Siciliano2, Mary Coolbaugh-Murphy2,5

Departments of 1Molecular Genetic Technology Program, 2Genetics, 3Biomathematics, 4Gastrointestinal Medicine and Nutrition, and 5Epidemiology; The University of Texas MD Anderson Cancer Center, Houston, TX

Abstract. Microsatellites are short tandem repeats of deoxyribonucleic acid (DNA) sequences which are distributed throughout the genome. Tumors in patients with Lynch syndrome tend to accumulate mutations in microsatellites at a much higher rate than other sequences in the genome resulting in microsatellite instability (MSI). This is due to germline mutations in mismatch repair (MMR) genes. Using small pool-polymerase chain reaction (SP-PCR), previous studies have shown that mutant alleles can be detected in microsatellites of DNA from peripheral blood lymphocytes (PBLs) of Lynch syndrome patients at frequencies that were low, but significantly higher than frequencies in PBLs of age-matched non-Lynch syndrome controls. In the present study, SP-PCR detection of frequency of mutant MSI alleles (FMMA) was performed on PBLs and saliva samples from four sets of families. Each family set consisted of a mutation carrying affected proband (initial tumor bearer), a germline mutation-carrier sibling without tumors, and an age-matched normal control, either related (for 3 family sets) without mutation carrier status or unrelated (for 1 family set) without mutation carrier status. FMMA of saliva and PBL DNA were compared between each proband, sibling and control for each family set, and between family sets. In all five statistically significant saliva comparisons identified between germline mutation carriers (FMMA: 0.080–0.261) and normal controls (FMMA: 0.003–0.087), the measured FMMA were always higher in the carriers (p < 0.05). A logistic regression model of the data showed a significant increase in FMMA in saliva DNA from siblings with MMR mutation compared to the normal controls (p < 0.001). These results indicated that the increased FMMA observed in the saliva DNA as well as PBL DNA of MMR gene mutation carriers compared to normal controls are real and repeatable. Furthermore, the logistic regression also indicated that the FMMA seen in saliva were nearly double those seen in PBLs (p < 0.001). Saliva testing, a less-invasive procedure than PBL testing, is more sensitive and appears to be a viable alternative for identifying MSI in carriers with MMR mutations.

Key words: SP-PCR, MSI, saliva, MMR, HNPCC

Introduction

Microsatellites are short, repetitive deoxyribonucleic acid (DNA) sequences that during DNA replication are prone to mismatching errors, which are normally repaired by the mismatch repair (MMR) system. However, with defective MMR systems, microsatellite instability (MSI) can occur. MSI is characterized by mutated or silenced MMR genes that lead to an accumulation of unstable microsatellites [1, 2]. Such an accumulation in various growth factor, regulatory or repair genes may alter their gene functions to a decreased or nonfunctional state [3]. High levels of MSI in somatic cells, predominantly (approximately 90% of cases) caused by inherited mutations in either MLH1 (MIM# 120436) or MSH2 genes (MIM# 609309), have been observed in patients with hereditary

0091-7370/11/0400-321. © 2011 by the Association of Clinical Scientists, Inc.
non-polyposis colon cancer (HNPCC) (MIM 114500) [1, 2, 4-7]. Allelic variations of the microsatellites of at-risk individuals, who are suspected carriers of the mutant or impaired MMR genes, are currently being used as early indicators of tumor formation in HNPCC or field effects where the instability indicates a tumor may form [8]. Such screening could be used as a preventative measure among individuals with HNPCC [9]. In this study, the utility of using less-invasive method of collecting saliva as a surrogate for blood draw is explored for use in SP-PCR MSI analysis in at-risk individuals.

Currently, standard polymerase chain reaction (PCR) is used to detect allelic variation microsatellites in somatic cells [10]. However, the standard MSI PCR assay has limited sensitivity and requires that mutant fragments have a frequency of >0.2–0.25 for detection purposes [11]. Below that frequency level, the progenitor fragments saturate the reaction wells and overshadow the less-frequent presence of mutant fragments in standard PCR assays [11]. Thus, if current standard PCR assays can be modified to increase their sensitivity, earlier detection of individuals at risk for HNPCC may be possible.

In 2004 Coolbaugh-Murphy et al. developed one such alternative [11], a high-throughput modification of small pool-PCR (SP-PCR) that was both quantitative and more sensitive allowing for detection of FMMA levels <0.25 [11,12]. SP-PCR partitions DNA fragments into multiple small pools containing only single genome equivalent (g.e.) amounts of DNA, trapping progenitor and low-frequency mutant alleles into multiple small pools where infrequently occurring (frequency as low as 0.01) alleles can be amplified and thus identified and quantified without interference from the overwhelming presence of progenitor fragments [11]. This modified SP-PCR technique was used recently to accurately measure levels of MSI in peripheral blood lymphocytes (PBLs) from patients with HNPCC, carriers of MMR mutations, and patients with MSI-H (defined as having at least 2 of 5 microsatellite markers showing instability) sporadic colorectal cancer [11-14].

Analyzing MSI levels in DNA extracted from PBLs is a less-invasive method than is sampling directly from tumor tissues [15, 16]. Nevertheless, screening with PBLs is still considered relatively more invasive and inconvenient than is a collection method such as buccal swabbing. The current literature indicates that individuals with inherited germline MMR mutations may show MSI in various normal somatic tissues [15,16]. This raises the possibility that other tissue sources besides PBLs could be sampled in a less invasive fashion than a needle stick. Saliva is one such alternative for testing MSI in carriers of germline MMR mutations, and the ease of collection would enable testing of more individuals that are in remote or distant locations and may potentially increase participation.

The present study aims to address the feasibility of detecting FMMA by conducting SP-PCR analysis on cellular components of saliva as an alternative to assays using PBLs as the source of DNA for MSI screening in patients with HNPCC and carriers of MMR mutations. This study was motivated by previous reports of circulating lymphocytes detected in samples of human saliva along with the presence of buccal (epithelial) cells [17,18]. To test the feasibility of using saliva samples instead of PBLs, the MSI levels of three loci (D2S123, D5S346, and D17S518), the loci previously studied by Coolbaugh-Murphy et al. [13,14,19], in saliva and PBL samples were evaluated and compared.

Materials and Methods

Study Participant Selection. Material from individuals participating in the MDACC study came from HNPCC families registered in the MDACC Epidemiology Department database. Probands were identified retrospectively, then siblings were prospectively recruited. The probands met the Amsterdam criteria for identifying HNPCC families. As the initial tumor-bearing individual (proband) in a family was identified, their tumor samples were analyzed for MSI using standard PCR at the five loci suggested by the NCI criteria [20]. In addition, their tumor samples were sequenced to determine the type of mutations in the MLH1 or MSH2 gene (Table 1). In addition, any current or previous cancers of any type were recorded. PBL DNA from recruited siblings was subjected to sequencing to identify mutation carriers. In addition, proband and sibling DNAs were subjected to large-pool PCR of the MSI loci to establish their progenitor genotypes.
All HNPCC probands underwent genetic evaluation and/or clinical management at MDACC. The Epidemiology Department at UTMDACC has an extensive database for tracking patients. Most often, not only the patients but also their siblings or children have undergone evaluation at MDACC. An IRB approved this work with the patient data and sample set. The criteria chosen for obtaining the ideal family sets for this study included: an HNPCC proband with a known MMR mutation, an MMR mutation-carrying sibling without known cancer, and a sibling not carrying an MMR mutation and with no known cancer to serve as normal control. Of the families identified from the epidemiology database, ten families met these criteria and four of these families participated in this study. (Because of the small number of families who met our criteria for proband, mutation carrier sibling and non-mutation carrier sibling, the more conservative Amsterdam criteria were chosen to limit confounding factors in this study.)

Collection of Peripheral Blood and Saliva. Two sets of test samples, PBLs and cellular components of saliva, from each of the family sets were studied. PBLs were collected as the families came to UTMDACC for their yearly checkup. The normal control (NC) PBLs were either from a sibling shown not to have inherited the mutation or a random age matched donor from the UTMDACC Blood Bank. If the NC siblings were available, they were matched by gender as well as age +/- 5 years of the affected siblings when possible. If matched NC sibling samples were not available, then additional NC samples were matched from a set of 426 normal control PBLs obtained from the UTMDACC Blood Bank identified by gender (212 females and 214 males) and age (range from 18-67).

Cellular components of saliva samples were collected using Oragene kits (OG-250, DNA Genotek, Ontario, Canada) (http://www.dnagenotek.com). Prior to the mailing of the kits, a letter was sent first to the subjects assessing their interest in participating in this project. A personal phone call followed within two weeks by one of the patient recruiters from the Epidemiology Department at UTMDACC. If the patient agreed, then an IRB approved consent form along with the Oragene kit was mailed to them. Each sample was tested for MSI using the SP-PCR method. All NC saliva samples came from either a sibling or, in one case, a matched NC from the MDACC blood bank. (Table 1).

Table 1. HNPCC Probands and Germline MMR Mutation Carrier Siblings

<table>
<thead>
<tr>
<th>Family Set</th>
<th>Sample ID</th>
<th>Age at PBLs</th>
<th>Age at saliva</th>
<th>Gender</th>
<th>Gene</th>
<th>Mutation Description</th>
<th>Mutation Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS-1</td>
<td>605</td>
<td>49</td>
<td>50</td>
<td>F</td>
<td>MLH1</td>
<td>Intron 7, A to G @ 589</td>
<td>Splice/pathological</td>
</tr>
<tr>
<td>FS-1</td>
<td>606</td>
<td>44</td>
<td>50</td>
<td>F</td>
<td>MLH1</td>
<td>Intron 7, A to G @ 589</td>
<td>Splice/pathological</td>
</tr>
<tr>
<td>FS-2</td>
<td>607</td>
<td>39</td>
<td>42</td>
<td>M</td>
<td>MSH2</td>
<td>Ex 12 AG to AA @ 1760</td>
<td>Splice</td>
</tr>
<tr>
<td>FS-3</td>
<td>609</td>
<td>40</td>
<td>43</td>
<td>M</td>
<td>MSH2</td>
<td>Ex 12 AG to AA @ 1760</td>
<td>Splice</td>
</tr>
<tr>
<td>FS-4</td>
<td>611</td>
<td>63</td>
<td>65</td>
<td>F</td>
<td>MLH1</td>
<td>Intron 7 A to G @ 582</td>
<td>Splice/pathological</td>
</tr>
<tr>
<td>FS-4</td>
<td>612</td>
<td>61</td>
<td>62</td>
<td>M</td>
<td>MSH2</td>
<td>Intron 7 A to G @ 582</td>
<td>Splice/pathological</td>
</tr>
<tr>
<td>FS-4</td>
<td>613</td>
<td>59</td>
<td>61</td>
<td>F</td>
<td>MSH2</td>
<td>Intron 14, G to A @ 2209</td>
<td>Splice</td>
</tr>
<tr>
<td>FS-4</td>
<td>614</td>
<td>56</td>
<td>60</td>
<td>F</td>
<td>MSH2</td>
<td>Intron 14, G to A @ 2209</td>
<td>Splice</td>
</tr>
</tbody>
</table>

Abbreviations: Family set (FS), mutL homolog 1 (MLH1), mutS homolog 2 (MSH2).
to approximate single diploid genome levels. The initial estimated genome equivalent was derived from serial dilutions of the original concentrated stock of DNA. A final concentration of 6 pg/µl, which is equivalent to 1 g.e. or two DNA fragments per reaction well, was obtained (Figure 1).

Approximately 30,000 hemi-nested SP-PCR reactions were conducted on the diluted DNA of blood and saliva samples from 12 subjects. There were 3 samples of blood and saliva from four HNPCC families: one normal control, one mutation carrying unaffected sibling, and one mutation carrying, affected proband. For each DNA sample, approximately 100 alleles for each locus were distributed across 112 PCR replicates. Three MWG Biotech Primus-HT 384 Thermocyclers (High Point, NC, USA) were used to amplify the outer PCR products, which contained ~1 g.e. per well. Amplicons from the outer PCR were diluted then distributed using the Qiagen BioRapidPlate and Twister I robots (Valencia, CA, USA) to set up the inner PCR. Three additional Primus-HT 384 thermocyclers were used for the inner PCR. Then RapidPlate was used to multiplex all three loci’s 384-well trays for each sample, resulting in 384 wells containing SP-PCR products of multiple loci from the same sample. The resulting plates were submitted to UTMDACC’s DNA Core Facility for analysis on the ABI 310 Genetic Analyzer.

**Statistical Analysis.** For data analysis, ABI software was used (GeneScan) to score every small pool’s fragments on the basis of chromatograms of raw data. For each locus, the number of variants and progenitor alleles was tabulated. The distribution of alleles in the small pools followed a Poisson distribution model developed by Coolbaugh-Murphy et al. [19]. The Poisson distribution of fragments at ~ 1 g.e. indicated that 25-33% of the wells would be empty: some wells might have no fragments while others might contain as many as four. The estimated number of g.e. reactions and the numbers of each progenitor or variant were entered into SPPCR2.0 (http://www.hkasoftware.com/sppcr.shtml) to calculate the actual estimated g.e. to be used during the full 112-replicate run and reconfigured into a 384-well plate for high throughput methodology. The Poisson approximation was used to calculate the statistical significance of differences between the groups using their p values. Using asymptotic normality of the estimates, the estimated frequency divided by its standard deviation was approximately Gaussian. The approximation approached perfection as the number of small pools increased. While the Poisson approximation was used to examine the significance of difference in frequencies between groups, logistic regression was used to examine the statistical significance between sample types. From the logistic regression model, coefficients for each of the three sample types were derived and FMMAs were calculated by adding the mutant fragments seen in one sample group over all the fragments seen in that sample group. Further analysis of the samples was conducted by separating both blood and saliva fragments using the same model. The likely p value was calculated after comparing each of the mutation carriers (carrier siblings, probands) with the normal controls. A third model, a saturated model, was used to combine all possible combinations of factors to derive a better fit.

**Results**

Each family set consisted of a proband, a germline MMR mutation carrier sibling without a known
Table 2. Frequencies of MSI in three microsatellite loci in the PBLs and saliva of germline mutation carriers and normal controls.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Carriers</th>
<th>Normal Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>5000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>5001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>5002</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

---

**MSI in Saliva from HNPCC patients and germline MMR mutation siblings**

1. **Table 2.** Frequencies of MSI in three microsatellite loci in the PBLs and saliva of germline mutation carriers and normal controls.
history of cancer, and a negative control consisting of a sibling with no MMR mutations (when available) or a normal individual with no known cancer. We measured MSI in a total of 36 sets of loci (12 loci sets from normal individuals [4 subjects x 3 loci = one loci set], 12 loci sets from probands, and 12 loci sets from MMR mutation carriers) from PBL samples and a total of 34 sets of loci (12 loci from normal individuals, 12 loci from probands and 10 loci from MMR mutation carriers) from saliva samples (Table 2). In one sibling sample carrying a germline MMR mutation, the salivary DNA was not measurable by SP-PCR at 2 loci (D5S346 and D17S518).

Each patient’s data for three loci were combined for a three-locus average that we used to compare FMMA comparisons: germline MMR-mutation carriers vs. controls in both PBLs and saliva (p < 0.05), germline MMR-mutation carriers’ saliva vs. carriers’ blood (p < 0.05), and probands vs. controls in both PBLs and saliva (p < 0.01). Of the 5 statistically significant saliva comparisons between germline mutation carriers (FMMA from 0.080 – 0.261) and normal controls (FMMA from 0.003 – 0.087), all were cases where the FMMA were higher in the carriers (Table 2).

Logistic regression of the data showed a gradual progressive increase in FMMA from normal controls to MMR mutation-carrier siblings to probands in saliva and PBLs (Figure 2). Logistic regression analysis also showed that FMMA among the three categories (NC, HNPCC carrier, and proband) were significantly different from each other (p < 0.01). Implementing a saturated model by adding the effect of NC, proband, and carrier sibling with other factors (PBLs, saliva) and fragments seen at each of the three loci further improved the fit (p < 0.001). Our results also show the overall FMMA (0.0417) in saliva as being twice the FMMA (0.021) in PBLs (p < 0.001) within the four family sets included in our study. The FMMA between the two sample types (saliva & PBL) within the NCs were not significantly different (p > 0.05).

Discussion

Researchers have been analyzing DNA isolated from saliva using various molecular techniques such as whole genome amplification, PCR of restriction fragment length polymorphisms, Southern
MSI in Saliva from HNPCC patients and germline MMR mutation siblings

MSI in saliva of probands and carriers was significantly higher than detected in normal controls. Our present study demonstrated that variant microsatellite fragments can be detected in saliva using SP-PCR. Our collective data suggests that MSI in probands as well as MMR-mutation carriers can be measured with DNA from saliva at levels significantly greater than that of the normal background controls (p < 0.05). These data suggest that saliva DNA can be used as a less-invasive surrogate for PBL DNA. This could direct screening for rare variants such as early detection of MSI in HNPCC families.

Comparison between the MSI levels found in both saliva and PBLs revealed that the FMMA in the saliva were nearly twice those seen in the PBLs.
Several potential causes exist for this unexpected result, beginning with the age of the patient when the sample was collected. The PBL samples had been collected from patients when they were 63, 59, and 56 years, but their ages when saliva samples were collected were 65, 61, and 59 years.

Coolbaugh-Murphy et al. [19] showed that FMMAs increases with age, specifically from a FMMA of 0.019 in the 35–50 years age group to a FMMA of 0.034 in the 60–70 years age group [19]. However, the time differences between our samples ranged from 1 to 3 years. Thus, it is highly unlikely that a two-fold increase in FMMAs from saliva over FMMAs from PBL was caused by the difference in time of sample collection. Another possible explanation could be that buccal cells are a better surrogate for colonic cells than PBLs since both buccal cells and colonic cells are derived from ectodermal origin during the second week of embryonic development. In contrast, PBL is derived from mesoderm via an invagination of the ectodermal cells during the third week of fetal development [28].

MSI that is higher in saliva than it is in PBLs could also be attributed potentially to recent somatic cell mutation events in the oral and pharyngeal area of subjects. MSI levels have been shown to be increased in saliva from patients with oral, head and neck, and pharyngeal cancers [29-31]. Furthermore, factors such as smoking, alcohol consumption, nutritional deficiencies, and DNA tumor viruses can all contribute to cellular DNA damage that can result in elevated MSI levels in saliva [32,33]. Since the three loci in our study showed base-line MSI levels in the normal controls and remained at background levels for both saliva and PBLs, it seems highly unlikely that these sources alone caused significantly higher FMMAs in saliva samples than those in PBL samples. Nevertheless, the precise influence of environmental factors or other genetic modifiers resulting in these various phenotypes remains to be determined. To resolve this outstanding issue, future studies with larger sample sizes and with consideration for the aforementioned factors should be conducted. Finally, the different cell types themselves could cause the observed difference between the FMMAs. The cell type with higher proliferation, if it contained a germline MMR mutation, would logically carry the higher mutation rates [34]. Thus, the higher MSI levels seen in saliva than were seen in PBLs could be caused by the proliferation status of these two cell types in carriers of germline MMR mutations.

Using SP-PCR MSI detection to identify MMR mutation carriers at a pre-neoplastic stage and enrolling them into preventative programs could be another tool for the management of colon cancer patients or family members [27]. The stability of the sample collection kits that are available on the market today gives a broader outreach to patients or at-risk family members alike in communities where health care might not be as easily attainable as in some of the larger more populous areas.

**Limitations.** Despite the advantages of SP-PCR using cellular components of saliva samples, our findings indicate that its quantitative capability is limited. FMMA calculations in our current study likely yielded underestimates, especially for heterozygous alleles that were one repeat unit apart; a single repeat unit deletion mutant of the larger progenitor would be indistinguishable from the smaller progenitor and in the opposite case (a contramutant, where the smaller allele is mutated to the same size as the larger allele) would not be scored as a mutant [11-13]. Even though we may have underestimated FMMAs in our study, we still found that SP-PCR can detect an increase in FMMAs in cellular components of saliva samples above that in the normal population. Regardless if this increased FMMA is due to DNA repair haploinsufficiency, circulating tumor cells, or a field effect, the mere presence of an increased number of variants should serve as a red-flag to care-givers of HNPCC patients or family members.

In addition, SP-PCR may also underestimate FMMAs when it is impossible to differentiate mutants of the same size that originate from mutational events different from those being trapped and replicated more than once during the reactions [11,13,14]. Finally, SP-PCR does not provide significant insight into mutation rates [11,13,14].

“The current version of the SP-PCR method is labor intensive and limited and not intended as a clinical test to diagnose patients with HNPCC [14].” The purpose of our study was to show that elevated MSI can be detected in saliva samples, and
our results indicate that saliva samples could be used as a less-invasive alternative to PBLs in the detection of MSI in HNPCC patients. Additionally, our study confirms that carriers of germline MMR mutations exhibit significant MSI levels in constitutive tissues, including cellular material found in saliva samples. That such a phenomenon is present not just in PBL, as shown previously, but also in saliva, and at higher levels, further supports such an idea.

Summary and future directions. In summary, the results of our present study indicate that MSI in carriers of germline MMR mutations can be detected in saliva cellular material DNA at levels significantly higher than in normal controls and with a higher sensitivity than in PBLs. Further investigations with larger sample sizes are needed to establish better baseline levels of MSI found in the normal population’s saliva. The effect of separating buccal cells from lymphocytes in the saliva should also be explored. The absence of buccal cells in peripheral blood indicates that measuring MSI levels from each population of those two cell types may provide additional insight into the relationship between MSI and specific tissues, i.e. similar tissues derived from the same embryonic layer may exhibit similar MSI phenomenon and therefore are a better surrogate tissue for MSI measurement than PBL. In addition, the use of sensitive assays such as SP-PCR in detecting low levels of mutants could be incorporated as part of the routine colonoscopy examination whereby colonic tissues may be swabbed for testing if no polyps were identified, or alternatively, test exfoliated colon cells in stool. Additional studies are needed to refine early MSI testing methods in patients with HNPCC, and such studies would play an instrumental role in developing early detection and prevention programs to increase the rate of survival of patients with HNPCC or other forms of colon cancer.

Acknowledgement

This work received support from the National Cancer Institute (grants CA112508 to M.J.S, CA070759 to M.L.F., CA57730-R25T cancer prevention postdoctoral fellowship to M.C.M., and CA-16672 NIH Cancer Center Support Grant [DNA Analysis Facility and Biospecimen Extraction Resource] to J. Mendelsohn, TexGen). We received partial support from Dr. Marsha Frazier’s DNA extraction laboratory and her HNPCC Registry, as well as from Ms. Erika Thompson’s DNA Core Laboratory.

References


17. Theidal C, Prange-Krex G, Freiberg-Richter J, Bornhauser M, Ehninger G. Buccal swabs but not mouthwash samples can be used to obtain pretransplant DNA fingerprints from recipients of allogeneic bone marrow transplant recipients. Bone Marrow Transplant 2000;25:575–577.


