Elevated Microsatellite Alterations at Selected Tetranucleotides (EMAST) Is Not Attributed to MSH3 Loss in Stage I-III Colon cancer: An Automated, Digitalized Assessment by Immunohistochemistry of Whole Slides and Hot Spots

Martin M Watson*,†,‡, Dordi Lea*,‡,§, Hanne R. Hagland*¶ and Kjetil Soreide*,†,‡

*Gastrointestinal Translational Research Unit, Laboratory for Molecular Medicine, Stavanger University Hospital, Gerd Ragna Bloch Thorsens gate 8, 4011, Stavanger, Norway; †Department of Gastrointestinal Surgery, Stavanger University Hospital, Gerd Ragna Bloch Thorsens gate 8, 4011, Stavanger, Norway; ‡Department of Clinical Medicine, University of Bergen, Bergen, Norway; §Department of Pathology, Stavanger University Hospital, Gerd Ragna Bloch Thorsens gate 8, 4011, Stavanger, Norway; ¶Department of Chemistry, Bioscience and Environmental Engineering, University of Stavanger, Norway

Abstract

INTRODUCTION: EMAST is a poorly understood form of microsatellite instability (MSI) in colorectal cancer (CRC) for which loss of MSH3 has been proposed as the underlying mechanism, based on experimental studies. We aimed to evaluate whether MSH3 loss is associated with EMAST in CRC. METHODS: A consecutive cohort of patients with stage I-III CRC. Digital image analysis using heatmap-derived hot spots investigated MSH3 expression by immunohistochemistry. Fragment analysis of multiplex PCR was used to assess MSI and EMAST, and results cross-examined with MSH3 protein expression. RESULTS: Of 152 patients, EMAST was found in 50 (33%) and exclusively in the colon. Most EMAST-positive cancers had instability at all 5 markers, and EMAST overlapped with MSI-H in 42/50 cases (84%). The most frequently altered tetranucleotide markers were D8S321 (38.2% of tumors) and D20S82 (34.4%). Subjective evaluation of MSH3 expression by IHC in tumor found ≤10% negative tumor cells in all samples, most being ≤5% negative. Digital analysis improved the detection but showed a similar spread of MSH3 loss (range 0.1–15.7%, mean 2.2%). Hotspot MSH3 negativity ranged between 0.1 to 95.0%, (mean 8.6%) with significant correlation with the whole slide analysis (Spearman’s rho = 0.677 P < .001). Loss of MSH3 expression did not correlate with EMAST. CONCLUSIONS: In a well-defined cohort of patients with CRC, loss of MSH3 was not associated with EMAST. Further investigation into the mechanisms leading to EMAST in CRC is needed.

Introduction

Microsatellite instability (MSI) is caused by defects in the mismatch-repair (MMR) family of proteins [1]. This results in mosaic populations of cells bearing microsatellite loci with diverse numbers of repeats due to uncorrected slippages during DNA replication. Silent or deleterious consequences arise according to the microsatellites affected and their location within the genome. In colorectal cancer (CRC), MSI is recognized as an alternative carcinogenic pathway to the chromosomal instability model, with a series of clinical and pathological implications [2]. MSI continues to be debated as a prognostic factor in CRC [3–5], and is implicated

© 2019 The Authors. Published by Elsevier Inc. on behalf of Neoplasia Press, Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1936-5233/19
https://doi.org/10.1016/j.tranon.2019.08.009
in the “hypermutated” or “immunogenic” consensus molecular subtype [6].

Elevated microsatellite alterations at selected tetranculeotides (EMAST) is a variant of MSI described in lung, skin, prostate, and other cancers, including CRC [7]. While MSI was initially defined as instability at mono- and dinucleotide repeats (e.g. CAn) [8], today commonly measured in a panel consisting exclusively of mono-nucleotides [9], the definition of EMAST is based on instability found in tetranculeotides (e.g. AAAGn).

MSI in CRC commonly displays loss of expression of MMR proteins such as MLH1, MSH2, MSH6 and PMS2. The MSH2 member of the MMR family can dimerize with either MSH6 or MSH3 to form the MutSδ or MutSβ complexes [1]. The latter is believed to have a higher affinity for repair of longer IDLs and mismatched sequences occurring during replication, such as tetranculeotides. MSH3 has therefore been implicated as a potential candidate to explain instability at longer microsatellites, as found in EMAST.

In vitro, MSH3 dysfunction was associated to instability at several tetranculeotide loci in MLH1- and MSH3-deficient CRC cell lines via whole chromosome transfer, as well as silencing/knockdown studies [10–12]. Additionally, it has been suggested that activity of MSH3 could be impaired by its dislocation from the nucleus to the cytosol, a process possibly mediated by interleukin-6 in a context of oxidative stress in CRC cell lines [12,13]. Furthermore, the cancer genome atlas (TCGA) consortium described MSH3 frameshift mutations—and not point mutations—as common (40%) in a subclass of CRCs defined as hypermutated and microsatellite-unstable [14]. Later, it was shown how MSH3, specifically in CRC, represents a frequent target of frameshift mutation, as opposed to the promoter hypermethylation that occurs at MLH1 in MSI CRCs [15]. The fact that the MSH3 gene contains a mononucleotide-repeat locus could suggest that frameshift mutations in MSH3 are a consequence of instability at mononucleotides initiated by loss of MLH1. In the mentioned studies it was not reported whether the frameshift mutations found in MSH3 were silent or non-silent, and their effect on functionality of the protein can therefore not be inferred. Should MSH3 be proven as the biological driver of EMAST, a causal relationship between MSI and EMAST could therefore be speculated. Thus, the relationship between MSH3 and EMAST need to be investigated in clinical cohorts. However, to date only 3 studies in relationship between MSI and EMAST could therefore be speculated.

Materials and Methods

The patient cohort was derived from the ACROBATICC project [18] (clinicaltrials.gov identifier: NCT01762813) and is conducted in accordance to national regulations and approved by regional ethics committee (REK Helse Vest, #2012/742). Written informed consent was obtained from each participant prior to inclusion in the study.

Patient Material

Formalin-fixed, paraffin-embedded (FFPE) tumor and normal tissue derived from stage I–III surgically removed CRC was used in this study. Appropriate slides were assessed by a certified pathologist and representative tissue blocks selected for DNA extraction, fragment analysis and immunohistochemistry.

EMAST and MSI Analyses

FFPE blocks were selected by an experienced pathologist and 4 × 10 μm sections were cut at a microtome. Automated DNA extraction was carried out using AllPrep DNA/RNA FFPE kit (Qiagen, Hilden, Germany) on a QiaCube instrument (Qiagen) according to manufacturer's instructions. Nucleic acid concentration and purity were measured on a NanoDrop 2000 (ThermoFischer scientific, Waltham, USA).

Multiplex PCR reactions (one for each MSI and EMAST) were set up for tumor and normal DNA from each patient. TypeIT microsatellite (Qiagen) master mix, together with a blending of 5 × 5′-fluorescently labeled primer pairs was used for each reaction. PCR conditions were as follows: 5′ at 95 °C (initial denaturation and enzyme activation), followed by 37 cycles of 30′ at 95 °C (denaturation), 90′ at 55 (MSI) or 57 °C (EMAST, annealing) and 30′ at 72 °C (extension). A final extension step for 30′ at 60 °C. The primers for EMAST were specific to the tetranculeotide loci MYCL1, D20S85, D20S82, D9S242 and D8S321 [19]. The primers for MSI were specific for BAT-26, NR-21, NR-24 and NR-27 [9,20], which are all quasimonomorphic mono-nucleotide repeats with a high fidelity to high-frequency MSI (MSI-H) as shown previously [21]. To define a tumor as EMAST and/or MSI-H, at least 2/5 markers needed to be unstable in their respective panels.

MSH3 Immunohistochemistry

Antigen retrieval and antibody dilution were optimized prior to the study onset. From FFPE blocks, 2 μm sections were cut and mounted onto Superfrost Plus slides (Menzel, Braunschweig, Germany). The sections were incubated at 60 °C for 1 h and then placed in the Dako Omnis autostainer (DAKO Agilent, Santa Clara, CA, USA). Automated protocol from the manufacturer was followed. Following deparaffinization and rehydration, antigen retrieval was performed at 97 °C for 30 minutes, and the slides were then incubated with the primary anti-MSH3 antibody (rabbit monoclonal anti-human MSH3; AbCam, Cambridge UK), clone EPR4334 (2), diluted 1:100 for 1 h. A peroxidase-DAB detection kit (Envision +, DAKO) was used to visualize the immune-complex. Sections were then counterstained with hematoxylin, dehydrated in increasing concentrations of ethanol and mounted manually.

Subjective IHC Score

Slides were evaluated and scored by an experienced pathologist for nuclear positivity of MSH3 (given as per cent, %) blinded to MSI and EMAST status of each case. A composite high-resolution image at 20× magnification of each slide was obtained with a Leica SCN400 scanner and uploaded onto an internal digital image hub for image analysis.

Digital Image Analysis

To increase scoring sensitivity of MSH3 expression, digitalized whole-slide and hotspots scoring of positive–negative nuclei in the tumor portion of the sections was performed with the aid of Visiopharm Integrator System software (VIS; Visiopharm A/S, Hoersholm, Denmark). An image analysis algorithm using Bayesian classification methods was built in an app-based tool which allowed for identification of tumor tissue within the scanned slides, and for the highlighting of its contours (Figure 1, A–C). Manual revision of each slide ensured then rigorous exclusion of tissue folds, stroma, necrotic areas, immune and
blood cells, and normal tissue from the analysis. A second, app-based algorithm was then developed to allow for the marking of positive and negative cells with colored labels (Figure 1, D–K). A heatmap based on the label associated with MSH3-negativity was created for each of the whole slides (Figure 1, E–F), for unbiased placement of one 0.8 mm² hotspots on the areas where the concentration of MSH3-negative cells was highest within the slides (Figure 1, G–H). Relative MSH3 negativity (both whole-slide and hotspots) was then derived from the ratio of negative label area and total negative and positive label areas, as calculated via the developed classifier.

Statistical Analysis
All statistical analyses were carried out using IBM SPSS statistics v. 25. Chi-square or Fishers Exact test were used for categorical variables. Relationship between different operators (pathologist/digitalized whole-slide/digitalized hotspot) in the scoring of MSH3 expression were tested using the Spearman’s rank order correlation. All tests were two-tailed with statistical significance set at $P < .050$.

Results

Patients Characteristics
Median age was 71.5 years (range 37.0–92.0), female patients were 85 (56%). There were 31 (20%) rectum and 121 (80%) colon cancers, 71 (59%) of which were in the proximal tract.

Tumor stages were equally represented with 51, 51 (34%) and 50 (33%) cases for Stage I, II and III, respectively.

EMAST and MSI Analysis
EMAST was present in 50 (33%, Figure 2, A and B) and MSI-H in 44 (28%) of 152 tumors, all of which in the colon (none in the rectum) and 90% in the proximal part of the colon. EMAST was positively associated with MSI (42/50 EMAST we also MSI-H, 84% overlap; $P < .001$), but not with tumor stage.

Almost half of the cohort (45%) showed no instability at any EMAST marker, while most EMAST-positive patients had instability at all 5 markers (Figure 2, C and D). The most frequently mutated marker was D8S321 (38% of tumors), followed by D20S82 (34%). The marker with the least events in microsatellite-stable tumors (most specific to EMAST status) was D20S85 (Table 1). Thirty-four (34) tumors had one unstable EMAST marker and were thus classified as EMAST-negative (Figure 2D).

MSH3 Analysis
Whole slide microscopical evaluation of tumor area for MSH3 loss found only <10% of the tumor nuclei with negative stain in all samples, most being ≤5% negative (Figure 3, A and B).

Digital analysis (Figure 3) showed a similar spread (range 0.1–15.7%, mean 2.2%). Hotspot MSH3 negativity ranged between 0.1 to 95.0%, (mean 8.6%) with a significant correlation between the two sets of measurements (Spearman’s rho $= 0.677 P < .001$), indicating that the measurements in the hotspots are indicative of the rest of the tumor.

MSH3 loss in hotspots was categorized in subclasses according to different cut-off points (1%, 5%, 10%, 25% and 50%) to establish whether a certain degree of protein loss could account for EMAST presence. None of the subclasses correlated with EMAST status, nor MSI-H or disease recurrence.

Heterogeneous expression of MSH3 was noted in some cases, with nuclei staining only partially positive for the protein (Figure 3C). However, these findings were not related to EMAST nor MSI.

Discussion
Based on a robust, automated, and digitalized protocol of IHC assessment with a verified MSH3 antibody, this study could not demonstrate an association between loss of MSH3 and EMAST in CRC. While other mechanistic contributions of MSH3 to EMAST

Figure 1. MSH3 immunohistochemistry virtual image analysis process. (A–C) An initial gross exclusion of stroma, tissue folds and normal tissue is carried out, selecting a “work area” where an app-based algorithm is then run to specifically mark tumor cells and exclude stroma. This results in a highlighted region of interest (ROI, marked in blue); (D) A second app-based algorithm classifies cells on the basis of their positivity (green) or negativity (blue) for the MSH3 protein; (E–F) A heatmap is created to highlight areas on the whole slide where the highest concentration of negative cells are located; (G) Based on the heatmap created in (E), a 0.8 mm² round ROI (hotspot, light blue circle) is placed and becomes the focus of the analysis; (H) Fully classified, hotspot-derived ROI.
cannot be ruled out based on the current experiment, this study suggests that neither protein loss, nor protein translocation (e.g. from cytosol to nuclei) is likely to be the cause of EMAST in patients with CRC. Thus, other mechanisms to EMAST must be considered beyond the notion that loss of MSH3 is essential to EMAST development [22]. Several points warrant further discussion.

The role of MSH3 expression in EMAST is debated and not yet fully understood. Some previous mechanistic investigations have based their experimental studies on effects seen in cancer cell lines [12,13]. In human CRC, however, one study showed high degree of MSH3 loss found in EMAST-positive cancers [16], but others have found no significant association between MSH3 loss and EMAST [17].

One possible explanation for the discordance between the current findings and previous results could be the source and type of antibody used for MSH3. One previous study used a clone to MSH3 that is no longer on the market [17], while neither the source nor clonality of the antibody used in a second study [16] could be reproduced. A third study used an affinity-purified rabbit polyclonal antibody [10]. The antibody used in the current study was developed and tested for human immunohistochemistry protocols, thus thoroughly validated using appropriate positive and negative controls.

**Figure 2. EMAST analysis.** Electropherograms of multiplex PCR fragment analysis are shown for (A) an EMAST-negative and (B) an EMAST-positive patient. Arrows indicate extra PCR products at +/− 4n bp in unstable markers. (C) Stacked bar population graph showing frequency of instability (red) at each EMAST marker, in EMAST-positive and EMAST-negative populations. (D) Bar chart showing proportions of patients grouped by total number of unstable EMAST markers. For patients bearing only 1 marker mutated (not EMAST, according to our thresholds) the bar is stacked to specify each marker’s abundance.

**Table 1.** EMAST markers, genomic loci (GRCh38/hg38 assembly), repeat type (− strand), and frequency of mutation.

<table>
<thead>
<tr>
<th>MARKER</th>
<th>LOCUS</th>
<th>REPEAT</th>
<th>EMAST+</th>
<th>EMAST−</th>
<th>MSI</th>
<th>MSS</th>
<th>all tumors</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYCL1*</td>
<td>1p34.3</td>
<td>AAAG</td>
<td>40 (88.9)</td>
<td>5 (11.1)</td>
<td>38 (84.4)</td>
<td>7 (15.6)</td>
<td>45 (29.8)</td>
</tr>
<tr>
<td>D9S242*</td>
<td>9q33.3</td>
<td>AAAG</td>
<td>39 (84.8)</td>
<td>7 (15.2)</td>
<td>38 (82.6)</td>
<td>8 (17.4)</td>
<td>46 (30.5)</td>
</tr>
<tr>
<td>D20S82*</td>
<td>20p12.3</td>
<td>AAAG</td>
<td>41 (78.8)</td>
<td>11 (21.2)</td>
<td>36 (69.2)</td>
<td>16 (30.8)</td>
<td>52 (34.4)</td>
</tr>
<tr>
<td>D8S321, n (%)</td>
<td>8q24.21</td>
<td>AAAG</td>
<td>48 (82.8)</td>
<td>10 (17.2)</td>
<td>40 (69.0)</td>
<td>18 (31.0)</td>
<td>58 (38.2)</td>
</tr>
<tr>
<td>D20S85, n (%)</td>
<td>20q12</td>
<td>AAAG</td>
<td>26 (96.3)</td>
<td>1 (3.7)</td>
<td>25 (92.6)</td>
<td>2 (7.4)</td>
<td>27 (17.8)</td>
</tr>
</tbody>
</table>

N = 151.
Subjective selection and cherry picking of regions of interest to measure IHC stains may be another issue across studies. In the current study, in order to account for variation and involuntary selection bias in the analysis of MSH3 expression, we constructed a heatmap-based hotspot analysis in addition to evaluation of whole slides. The hotspot approach is presently proving useful in the evaluation of prognostic biomarkers with thresholds of low expression, such as Ki67 in breast cancer [23], and has been implemented in the Norwegian, Danish and Swedish pathology guidelines [24–26]. However, results will still largely depend on the robustness of the antibodies used and their reproducibility. For MSH3, the previous studies used antibodies that are no longer on the market, thus difficult to reproduce.

Immunohistochemistry is a complex multi-step procedure, performed by laboratories on a range of instruments and interpreted subjectively by pathologists according to guidelines that are often designed for a specific diagnostic test. Autostainer instruments and digital image analysis have reduced the discrepancies originating from manual experiments [27], although this technique is still often performed manually in research labs and as such open to a multitude of sources of variation [28,29]. To reduce bias and misinterpretation of results to the best of our abilities, this study employed automated staining together with objective, automated assessment.

We found no association between MSH3 expression and EMAST in CRC in the present study. The findings may have implications beyond CRC as EMAST has been described in tumors other than CRC [7]. A recent study in patients with pancreatic cancer that described EMAST in up to 40% of patients found no inactivation of MSH3 in the tumors [30], which is in line with the current results, although in a different cancer type. Furthermore, MSH3 frameshift mutations were found to be specific for CRC, but not for endometrial cancer [15], which is also known for harboring high levels of EMAST [31]. These observations, taken together with the results of this study and the general disagreement in the literature, discourage the proposed role of MSH3 loss as a universal biological mechanism underlying EMAST.

The large overlap of MSI and EMAST in CRC found in the current study correlates well with results from the literature [17]. EMAST could represent an exacerbation of MSI rather than a separate occurrence. In the current study we found EMAST to occur in colon cancers and to be associated with MSI features, as described previously [19]. Shared features between MSI and EMAST include prevalence in the proximal part of the colon and, in the case of the present study, an absence or low prevalence in rectal cancers. Two reports [16,32] describe a high prevalence of EMAST among rectal cancer patients, however in one (a rectum-only study) EMAST prevalence was higher in African American than Caucasian patients [32]. In the second report, describing 61% of rectal cancers as EMAST, the cohort was based on patients of Asian ethnicity [16]. In contrast, the present cohort consists predominantly of Caucasian patients. Considering the documented variation of MSI-H across demographic factors such as gender and ethnicity [33], EMAST might well follow a similar pattern, thus explaining discrepancy in rectal distributions. The current results are nonetheless consistent with other studies [16,17,34] finding EMAST prevalently in colon, and more specifically in the proximal part of the colon.

In conclusion, the mechanism leading to genomic instability in tetranaucleotides and expressed as EMAST is still largely unknown at present. Indeed, EMAST may, as suggested for conventional MSI, be related to epigenetic mechanisms that occur with aging (e.g. epigenetic loss of DNA repair mechanisms). EMAST may thus represent an epiphenomenon of age and tumorigenesis rather than a specific tumor-driving trait per se. On the other hand, MSI-induced frameshifts could affect MSH3 functionality—and lead to EMAST—in ways that the antibody used in this study could not highlight. Further investigations into the mechanisms of EMAST is warranted.

References


