Left-Sided Colorectal Cancer Distinct in Indigenous African Patients Compared to Other Ethnic Groups in South Africa.

M. McCabe (michelle.mccabe@nhls.ac.za)
National Health Laboratory Service
C. Penny
University of the Witwatersrand
P. Magangane
University of the Witwatersrand
S. Mirza
University of the Witwatersrand
Y. Perner
National Health Laboratory Service

Research Article

Keywords: colorectal cancer, microsatellite instability, Chi-square tests, miRNA

Posted Date: October 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-845611/v1

Version of Record: A version of this preprint was published at BMC Cancer on October 24th, 2022. See the published version at https://doi.org/10.1186/s12885-022-10185-3.

Abstract

Introduction: A large proportion of indigenous African (IA) colorectal cancer (CRC) patients in South Africa are young (<50years), with no unique histopathological or molecular characteristics. Anatomical site as well as microsatellite instability (MSI) status have shown to be associated with different clinicopathological and molecular features. This study aimed to ascertain key histopathological and miRNA expression patterns in microsatellite stable (MSS) and low-frequency MSI (MSI-L) patients, to provide insight into the mechanism of the disease.

Methods: A retrospective cohort (2011-2015) of MSS/MSI-L CRC patient samples diagnosed at Charlotte Maxeke Johannesburg Academic Hospital was analyzed. Samples were categorized by site [Right colon cancer (RCC) versus left (LCC)], ethnicity [IA versus other ethnic groups (OEG)] and MSI status (MSI-L vs MSS). T-test, Fischer's exact and Chi-square tests were conducted. Additional miRNA expression profiling was performed on IA patient samples.

Results: IA patients with LCC demonstrated an increased prevalence in males, sigmoid colon, signet-ring-cell morphology, MSI-L with BAT25/26 marker instability and advanced disease association. MiRNA expression profiling revealed unique clustering, with dysregulation of let-7 and miRNA-125.

Conclusion: This study revealed distinct histopathological features for LCC, and suggests BAT25/26, miRNAs let-7a-5p and miRNA-125a/b-5p as negative prognostic markers in African CRC patients. Larger confirmatory studies are recommended.

Introduction

Right-sided colon cancer (RCC) and left-sided colon cancer (LCC) show distinct mechanisms of development and are associated with different clinicopathological features.[1–3] During embryological development, the RC develops from the midgut and the LC from the hindgut, supporting the view that RCC and LCC develop through different developmental/embryological pathways genetic mechanisms.[1,4]. The incidence of RCC (~30%) is lower than LCC (~70%). RCC presents with larger tumours, a higher rate of tumour node metastases (TNM stage), mucinous features, and comprises of an overall poorer survival than LCC.[1–7] Older female patients are at higher risk of developing RCC compared to younger male patients associated with increased risk of developing LCC.[1] The literature shows population groups with a lower risk of developing CRC moving to high-risk areas acquire the risks associated with the new area, and this could be linked to dietary, environmental, cultural and genetic factors.[1,8] This speaks to a possible role for the increasing incidence of CRC in indigenous African (IA) patients moving from rural to urbanized areas.[9–11]

There are 3 main pathways involved in the development of CRC: 1) Microsatellite instability (MSI) caused by a defective mismatch repair (MMR) system due to MLH1 promoter methylation silencing, 2) Chromosome instability (CIN) pathway which develops due to gross chromosomal changes and 3) the CpG island methylator phenotype (CIMP) pathway, arising through methylation of CpG islands in promoter sequences, leading to inactivation of tumour suppressor genes throughout the genome.[12,13] MSI and CIMP tumours mostly occur in the right colon, whereas CIN CRC is associated with LCC.[12,14,15] Four main consensus molecular subtypes (CMS) were established in 2015. These are differentiated by unique molecular features: CMS1 (14%, MSI pathway, immune activation); CMS2 (37%, Canonical WNT/MYC pathway, epithelial signature); CMS3 (13%, epithelial and metabolic dysregulation), and CMS4 (23%; Mesenchymal TGF- β pathway; stromal invasions and angiogenesis). [12,16]

To date, fewer research outputs on low frequency microsatellite instability (MSI-L) CRC have been published compared to MSI-H and MSS CRC. MSI-L is usually grouped with MSS CRC, as literature states all CRCs display some level of MSI.[17,18] Some researchers interpret MSI-L tumours as precursors of MSI-H CRC, whereas others believe it to be a

completely separate entity.[19,20] MSI-L tumours have illustrated different clinicopathological features and have been considered to be a worse prognostic group in a few CRC studies.[21–24] Jass et al. reported that MSI-L LCC showed distinctive clinicopathological features, with a male predilection, a moderately differentiated histopathological grade, KRAS mutations, CIMP-Low status and DNA aneuploidy. In contrast, MSI-L RCC was found to occur more frequently in females, being associated with a serrated adenoma precursor lesion, mucinous adenocarcinoma histological subtype and poorly differentiated grade. BRAF mutations, CIMP-High status and diploid DNA content, features associated with a worse prognosis, were also associated with the MSI-L RCC group.[25]

MiRNA expression profiling has been shown to differentiate between CMS subgroups, as well as sporadic versus hereditary MSI CRC. Distinct miRNA expression levels have been associated with survival outcomes.[26–30] The CMS2 subtype is associated with high expression levels of the mir-17-92 cluster. The CMS3 subtype is associated with low expression of let-7 family, and CMS4 subtype, with low expression of the mir-200 family.[30]

A disproportionate number of indigenous African (IA) patients display a younger age of onset (<50 years of age) with no distinct histopathological features to assist with early diagnosis and management.[31–34] Previous work by McCabe et al. described MSI-H CRC in detail according to ethnicity groups and found an increased association of MSH2/MSH6 MMR protein expression loss in right sided CRC in young IA patients.[35] This study aims to characterize proficient MMR (MSS/MSI-L) CRC, by ethnicity (IA versus OEG) and anatomical site (LCC versus RCC), with additional miRNA expression profiling analysis to potentially identify a unique subtype associated with young IA CRC patients.

Methodology

Patient demographics and tumour histopathological characterization:

This retrospective study was conducted on a 5-year cohort of 369 CRC patients with a proficient(p) MMR (MSS/MSI-L) status diagnosed at the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) branch of the National Health Laboratory Service (NHLS) between the period of 2011-2015. Informed consent was waived by the Human Research Ethics Committee (HREC) (Medical) of the University of the Witwatersrand for this study, as research was conducted under the institutional blanket ethics clearance (M10744) obtained from the HREC (Medical), which allows for research to be carried out on all archived pathology specimens without informed consent from study participants. Additional project-specific ethical clearance was also obtained from the HREC (Medical) (M120994), and all tests were performed according to the relevant guidelines and regulations. A total of 369 pMMR CRC cases were categorized into 4 groups, see Table 1. Tumour site was the first group, based on RCC (tumour primary site proximal to splenic flexure) versus LCC (tumour primary site distal from splenic flexure). Ethnicity was the next category, stratifying Indigenous African (IA) versus Other Ethnic Groups (OEG) [White (Caucasian), Coloured (Mixed Ancestry), Indian (Asian)] within the tumour site categories. Demographic and histopathological information were analyzed within these categories i.e., gender, age, tumour subtype, grade, site, TNM stage (American Joint Committee on Cancer [AJCC] TNM stage), presence of tumour infiltrating lymphocytes (TIL) using the Klintrup-Mäkinen scoring assessment, Crohn's-like inflammatory reaction (CLR), polyp subtype, venous, perineural and lymphatic invasion, all of which were obtained from histology reports. No family histories were available from these reports.

MSS versus MSI-L molecular subtyping:

Samples were screened for proficient MMR status through MMR immunohistochemistry (IHC) and/or MSI polymerase chain reaction (PCR). The MMR IHC panel included antibodies targeting MutL Homolog 1(MLH1), MutS Homolog 2 and 6 (MSH2/ MSH6) and Post Meiotic Segregation Homolog-2 (PMS2) protein expression. Only samples with a MMR

proficient profile detected via IHC and a MSS or MSI-L profile determined via PCR were included in this cohort.[35] MSI PCR included the 5-mononucleotide PCR panel (NR27, NR21, NR24, BAT25 and BAT26); and a MSS or MSI-L result was ascribed when an allelic size varied in none or only one of the 5 markers respectively[35,36]. The literature indicates that people of African descent exhibit normal variation within loci BAT25 and BAT26.[37–40] A multipopulation study by Buhard et al. 2006, revealed approximately 10% show normal variation in one of five markers and 2% in 2 markers.[39] This panel remained the assay of choice due to the additional 3 quasimonomorphic mononucleotide repeat markers found in Caucasian and African germ-line DNA, ensuring the panel is extremely sensitive in detecting somatic alterations in MSI-H tumours and distinguishing between MSS/MSI-L tumours. As recommended by the authors Suraweera et al. 2002, in tumour samples with instability in BAT25 and/or 26 markers with a proficient MMR profile via IHC, matched normal samples were assessed to establish the true instability status. In cases where these markers matched instability in normal colon tissue, the status was regarded as normal or germline variation an reported as MSS. In biopsies this was not possible due to the limited size and mixture of normal and neoplastic tissue, increasing the chance of contamination when assessing normal tissue.

Statistical analysis:

All data was collected in an excel sheet and statistical analysis was performed using Stata Intercooled 7.0 (Stata, College Station, TX, USA) and Graphpad Prism version 9.0 (Graphpad Software, La Jolla, CA, USA). Column and contingency statistical analysis was applied using t-tests, Fischer's exact and Chi-square tests; and a result with a P value less than 0.05 was considered statistically significant. Additional descriptive analysis of MSI status (MSI-L versus MSS CRC) stratified by tumour site (LCC versus RCC) was conducted to determine if any association occurred with demographic data (gender, age and ethnic groups) as well as TNM stage and Crohn's-like inflammatory reaction (CLR). In the instance of multiple variable analysis, the False Discovery Rate (FDR) p-value (q-value) was applied and considered statistically significant if $q<0.05$. (see Table 2).

MiRNA expression profiling:

A total of 15 random CRC samples from the IA patient cohort was categorized by molecular MSI subtype and tumour site into 4 subgroups (1. MSS-LCC 2. MSI-L LCC 3. MSS-RCC and 4. MSI-L RCC) (supplementary data table). Total RNA was extracted utilizing the Qiagen RNeasy® FFPE kit and reverse transcriptase polymerase chain reaction (RT-PCR) and miRNA expression profiling was carried out with the miScript® II RT kit and Human cancer pathways finder miRNA PCR Array: MIHS-102Z (Qiagen, Whitehead Scientific WHS), respectively. All tests were performed according to the manufacturer's instructions. Data was analysed using the Qiagen miRNA webtool (https://dataanalysis.qiagen.com/mirna/arrayanalysis.php), where MSS-LCC was used as the control group for comparative expression analysis. Differential miRNA expression is accepted as significant if the fold change in expression is greater than 2 or less than 0.5. A p-value however was not calculated as 3 biological replicates for each group was required for statistical analysis and was not feasible at the time of this study.

Results

Patient demographics. LCC revealed a higher frequency in males in IA than in OEG patients (64% vs 47%, respectively; P=0.0111) (Table 1). The IA patients were younger in comparison to the OEG patients (median age: 54 vs 62 years, respectively; P<0.0001).

Pathological characterization. Signet ring cell carcinomas (SRCC) were more frequently found in IA (14/213; 7%) versus OEG patients (2/147; 1%) (P=0.0221). When further stratified by site, SRCC was only associated with LCC in IA patients, as compared to OEG patients (P=0.0257).

MSI-L versus MSS molecular subtyping. Ethnicity was linked to mononucleotide instability markers, with BAT25 and BAT26 markers being more frequently unstable in IA patients (28/29; 97%). NR21, NR24 and N27 instability was commonly demonstrated in OEG patients (7/13; 54%) (P=0.0053). An increased rate of MSI-L CRC was found in the left colon, particularly in IA compared to OEG patients (21% vs 9%; P=0.0442). Within the left colon, the MSI-L subtype was associated with more advanced disease stage (III-IV) (8/9; 89%) as compared to MSS CRC (38/76; 50%) (P=0.0348) (Table2). CLR was associated with MSI-L LCC (6/9; 67%), when compared to MSI-L RCC (1/8; 12%) (P=0.00498), and MSS LCC (19/73; 26%) (P=0.0205) (Table 2). No further distinctions were found amongst ethnic groups and RCC, apart from young age (P=0.0103) (Table1).

miRNA expression profiling. IA CRC miRNA expression profiling is depicted by the non-supervised clustering of the entire dataset (Figure 1). Four distinct subtypes by MSI status and tumour site (1. MSS-LCC; 2. MSI-L-LCC, 3. MSI-L-RCC and 4. MSS-RCC) are illustrated. The clustergram grouped MSS-LCC and MSI-L LCC as separate molecular groups. In comparison, MSS-RCC and MSI-L RCC were more closely related. Two main miRNA family members (let-7 and miR-125) were substantially dysregulated in the 4 subgroups. MiR-125a was upregulated for MSI-L LCC, MSI-L RCC and MSS RCC by approximately 4.3K, 0.8K and 0.4K log fold, respectively. Let-7a was downregulated for MSI-L LCC and MSS RCC by -0.2K and -0.5K log fold, respectively. Isoforms let-7e and let-7f were also considerably downregulated in MSI-L LCC when compared to the other 3 subgroups. MiR-125b was downregulated in MSI-L LCC, MSS RCC and MSS LCC by approximately -5.6K, -2.6K, and -1.0K respectively, when compared to MSI-L RCC.

Discussion

CRC has been shown to have unique clinicopathological features associated with tumour site and different molecular subtypes (CMS 1-4).[30] CRC molecular subtypes and age of onset have also been described to vary considerably among geographically distinct ethnic groups. Within this cohort, (40%) of IA patients was shown to be younger (<50years) compared to OEGs with pMMR CRC. Patients of OEGs in this cohort displayed RCC with poor prognostic factors compared to LCC. Increased frequencies of HG tumours (11% vs 4%) advanced staged tumours (73% vs 49%), perineural invasion (22% vs 15%), mucinous and signet ring morphology (19% vs 4%) were seen in RCC. While the median age of onset was similar for LCC vs RCC (61 vs 62), more males however were diagnosed with RCC (59%) than LCC (47%).

Within the IA patient group, both left and right colon cancers showed similar frequencies for poor prognostic factors. Higher frequencies for HG tumours (11%), advanced tumour stage (62-64%), perineural invasion (24-26%), SRCC (6%), younger age onset (median age 54-55), with more males presenting with LCC compared to RCC (64% vs 54%).

When comparing ethnic groups and right versus left-sided CRC, significant differences were observed for IA patients with LCC. The IA population group showed a propensity to occur in males, within the sigmoid colon, to present with a SRCC histological pattern, and an MSI-L status. Notably, SRCC is recognized as a rare histological subtype (1%) of CRC and is associated with young adults in other geographical locations.[41] Previous studies have shown SRCC to have a RCC dominance. However, more recently, SRCC has been reported to have an even site distribution within the colon, with a slight male predominance.[42] Moreover, the SRCC histological subtype is known to have an adverse prognostic significance independent of tumour stage and molecular subtype.[43,44] Poor tumour grade and advanced TNM staging are usually associated with worse survival outcomes.[45,46] In this study, these features were found to have borderline significance in IA patients with LCC compared to OEGs, with slight increases in frequencies of HG (11% vs 4%, respectively; P=0.0930) and advanced disease stage (62% vs 49%, respectively; P=0.0922).

When assessing PCR confirmed cases only to accurately categorize MSS versus MSI-L CRC in right versus left CRC, significant associations were seen for MSI-L LCC with advanced disease stage and the IA ethnic group (Table 2). This data was perceived to be similar to the findings of Devaraj 2010, linking elevated microsatellite alterations at selected tetranucleotide repeats (EMAST) with advanced disease stage, rectal cancers and peri-tumoral infiltration in patients of African American (AA) descent.[47,48] Even though MSI-L status was not available for the data in the study of Devaraj et al. 2010, previously published data closely linked MSI-L tumours to EMAST and has been associated with a poorer prognosis.[21,23,49] The most frequent unstable markers in tumours from IA patients in our cohort were BAT25 (15/29; 52%) and BAT26 (13/29; 45%). MSI-L LCC tumours showed a tendency to be of a more advanced disease stage (AJCC TNM stage: III-IV) compared to MSS LCC (89% vs 50%, respectively; P=0.0348).

Previous studies have shown these markers to be polymorphic within the African population, linked to the theory that older population groups show increased genetic variation.[37–39] Within this cohort, 17 CRC patients had shown instability in one or both markers (BAT25 and/or 26) in tumour and matched normal tissue, with a proficient MMR protein expression profile. These samples were assumed to be due to increased genetic polymorphisms. What was interesting and noteworthy to mention, was all these patients were exclusively of IA descent, with no differences observed in allele deletion sizes (5-15bp deletion) between germline and somatic instability, and the majority were LCC (15/17; 88%) with advanced disease stage at diagnosis (7/10; 70%). Based on these findings, limiting as it is in size, BAT25/26 instability (whether of polymorphic/germline or somatic variation) was associated with advanced disease stage in proficient MMR LCC patients of IA descent. Even though survival data was not available for this study, literature has shown poor clinical prognosis and overall survival associated with MSI-L CRC, particularly for advanced disease stage CRC.[22–24,50] The somatic MSI-L LCC group was associated with CLR when compared to MSI-L RCC (67% vs 12%; P=0.0498) and MSS LCC (67% vs 26%; P=0.0205). Polymorphic/germline MSI-L tumours within 2 markers (11/17; 65%) however displayed no CLR. Literature has indicated tumours with CLR to have a better prognosis compared to stage-matched tumours without.[51–53] This raises a plausible argument that germline MSI-L tumours could have a worse prognosis compared to somatic MSI-L tumours, due to the lack of the host's immune response to the cancer.

BAT26 marker (26(A) repeats) is located in intron 5 of the MSH2 gene on chromosome 2p21. This marker is situated immediately downstream of exon 5, which is susceptible to large intragenic deletions and accounts for nearly a third of dMSH2 mutations.[54–56] Studies by Pastrello et al 2006 and Jaskowski et al 2007 indicated that instability in Bat26 was associated with overall instability of dMSH2 tumours. Confirmatory IHC to determine dMMR protein expression is therefore important, however exceptions of cases with mutations in intronic nucleotides close to splice sites could result in expressed non-functional proteins, such as (*MSH2* c.913G>A p.Ala305Thr) which has been reported with proficient MMR activity and a MSI-L genotype.[56] This variant however had no aberrant splicing and normal subcellular localization and interaction with MSH6 was shown.[56,57]

BAT25 (25(T) repeats) is situated within intron 16 of the c-kit proto-oncogene on chromosome 4q12. cKit (CD117) the receptor for Stem cell factor (SCF) involved in haemopoiesis has more recently shown to be involved in lymphopoiesis. The CD117 receptor has shown to be expressed on mature $CD8^+$ T cells following initial activation, suppressing differentiation and increasing its response to apoptosis. CD117 expressed $CD8^+$ T cells could therefore play a role in CD117-blockade, an important mechanism in tumour immune evasion. BAT25 instability in the CD117 gene could potentially play a role in immune evasion in MSI-L CRC, and additional studies are required to determine CLR and its association with MSI-L LCC.

A study by Carethers et al. 2014 showed an increased incidence of MSI-H CRC in AA patients, however had poorer prognosis and higher mortality rates compared to their Caucasian counterparts.[58] This was thought to be due to AA

patients showing a lower infiltration of CD8+ T cells compared to Caucasian patients, suggesting an altered immune function in AA patients. It has been well established that the increased tumour-infiltrating $CD4^+$ and $CD8^+$ T cells in patients with MSI-H CRC (due to the increased mutator phenotype of the tumour stimulating the host immune response) significantly improved patient outcomes when treated with immunotherapy when compared to MSS CRC with decreased immune response.[59,60] The study by Carethers et al. 2014 illustrate despite having the same disease subtype and stage, ethnicity can be a negative prognostic factor in CRC disease.

Previous miRNA expression studies in CRC have also identified distinct profiles among the different CMS groups. In this study, the clustergram demonstrated MSS LCC and MSI-L LCC to be distinct groups, whereas MSS-RCC and MSI-L-RCC were more closely related. MiRNA expression profiling of IA CRC samples identified two main miRNA family members, namely let-7 and miR-125, substantially dysregulated in the 4 pMMR groups stratified by molecular instability status and site: (1.MSS LCC; 2.MSI-L LCC; 3.MSS RCC; 4.MSI-L RCC). Let-7a-5p and miR-125b-5p were not expressed in the MSI-L LCC and MSS RCC tumours with more advanced TNM stage, compared to dominantly early TNM stage groups (MSI-L RCC and MSS LCC), which showed considerable upregulation. These results support the roles of miR-125b and let-7a as tumour suppressors in the development or progression of CRC in this cohort. MiR-125a-5p was however substantially upregulated in MSI-L LCC and MSS-RCC, indicating a possible oncogenic role of this miRNA in the development of these advanced stage groups. Well documented targets of let-7a-5p and mir-125b-5p in CRC are the oncogenes KRAS and HGMA2.[61–63] A study reported by Liu 2016 et al. revealed a negative correlation between let-7a-5p expression levels and overall disease-free survival, with down-regulated let-7a-5p expression and high HMGA2 protein expression. The patients demonstrated poorer survival rates compared to patients with high let-7a-5p expression.[64] This study also illustrated let-7a-5p expression levels in serum were positively correlated with levels found in tumour tissue, indicating let-7a-5p to be a potential clinical biomarker for prognostication in CRC patients.[64–66] Further, when upregulated miR-125b and let-7a levels were downregulated in cell lines and animal models and cell proliferation and invasion were inhibited, thus implying a possible therapeutic strategy for the treatment of CRC.[28,65–67] Limitations for this miRNA study were the lack of evaluation of comparable normal colonic tissue. Previous studies have illustrated that let-7 isoforms dominate in normal colon tissue, while in neoplastic tissue, expression levels of let-7 are reported to be significantly suppressed.[68–71] In addition, an OEG comparator group was not investigated, and only a small sample size evaluated, increasing the possibility of false positives or over exaggerated log fold expression levels.[72]

A frequency of 17% MSI-L CRC in SA CRC patients has been reported in our previous study.[35] Further evaluation in this study has demonstrated MSI-L LCC to occur predominantly in IA patients, and associated with advanced disease stage, with considerable number of germline/polymorphic MSI-L LCC also presenting at an advance stage compared to MSI-L RCC and MSS LCC. In view of the heatmap clustergram results, MSI-L LCC most likely represents a standalone subtype distinct from MSS CRC and is not a precursor of MSI-H CRC.[21,23,73–75]

Based on these findings, universal MSI PCR screening is recommended as a first-line screening method for all newly diagnosed CRC patients, to not only identify MSI-H CRC, but also increase the detection rate of MSI-L CRC. CRC is a heterogenous disease and more studies are required to unravel the complexity associated with the development of the disease by investigating different ethnic groups in the context of site. This study has shown that ethnicity and tumour site play an important role in the prognostication of tumours and should be taken into consideration for effective treatment planning, especially in geographical regions with diverse population groups such as South Africa. Limitations of this study include selection bias, as only samples with an MSI status were included. This resulted in smaller sample sizes for the analysis of certain features such as polyps, TILs, TNM staging, MSI-L and BAT25/26 instability status. The lack of universal screening for MSI within the study institution, as well as the inclusion of biopsy

samples in addition to resections, have contributed to providing limited information. Confirmation of MSI-L status in biopsies was not possible, increasing the likelihood of a small percentage of false positive MSI-L samples with normal variation.

A study by Ozaki *et al* found that MSI-L colon tissue occurred in a few but not all intestinal crypts, and both in malignant and normal tissues. The presence of MSI-L in non-neoplastic mucosa could indicate a primary step in tumorigenesis and could potentially be used as an early diagnostic and prognostic marker in CRC.[76]

Additional AA patient studies have illustrated increased frequencies of MSI-L/EMAST markers in rectal cancers most likely due to somatic inactivation of an alternative MMR gene (MSH3).[22,47,50,77–79] Dysfunctional MSH3 has shown to lead to MSI, appearing to be inflammation-related within the tumour microenvironment.[80] Regular intake of anti-inflammatory drugs such as aspirin and non-steroidal anti-inflammatory drugs (NSAID) has been reported to prevent the development of colorectal adenomas, tumour growth and progression, as well recurrence and metastasis after curative surgery, prolonging colorectal cancer patient survival.[81–87] Anti-inflammatories could therefore possibly have a positive effect not only in MSI-H CRC, but also in this subgroup of MSI-L LCC patients. In addition, due to the presence of CLR, MSI-L LCC could potentially be an eligible subgroup for immune checkpoint inhibitors in metastatic disease and further studies recommended.

Conclusion

This SA CRC study indicated that in considering categorization of CRC according to anatomical site, microsatellite instability status and ethnicity, unique clinicopathological features were identified. In particular, IA CRC patients with LCC are more likely to be male, have an MSI-L subtype, show BAT25/BAT26 marker instability and have advanced disease stage. This study suggests that BAT25/26 instability and the lack of expression of miRNAs let-7 and miRNA-125 are negative prognostic markers in African CRC patients, and larger confirmatory studies are recommended. Further exploratory studies of MSH3, EMAST, KRAS and immune cell infiltration in the tumour microenvironment are indicated in SA CRC patients. This will assist in establishing molecular profiles to accurately improve diagnostic, prognostic and personalized predictive markers for the effective management of early onset CRC.

Declarations

Compliance with ethical standards:

This was a retrospective study on already available biological material and ethical clearance was obtained from the Human Ethics Research Committee (Medical), University of the Witwatersrand Ethics clearance reference number: M120994. The need for informed consent was waived by the HREC (Medical), University of the Witwatersrand for this study, as research was conducted under the institutional blanket ethics clearance (M10744) obtained from the HREC (Medical), which allows for research to be carried out on all archived pathology specimens without informed consent from study participants. All experimental protocols were approved by the Postgraduate Research and Protocol committee within the Faculty of Health Sciences, University of the Witwatersrand, and all experiments were performed according to the relevant guidelines and regulations.

Data Availability:

The dataset generated and analysed during the current study are available from the corresponding author on reasonable request.

Consent for publication:

Not Applicable.

Conflict of interest:

The authors have no conflicts of interest to declare.

Funding

Griffin Trust Fund (GTCP016)

National Health Laboratory Services Trust Fund (94347)

Wits University FRC Fund (4550)

National Research Foundation (001.403)

Author Contributions:

MM, YP and CP were involved in the design and research methodology of the study. MM conducted the tests, performed the analysis and wrote the manuscript. PM contributed to the miRNA methodology and analysis. All authors reviewed the manuscript.

Acknowledgements:

Not Applicable.

Corresponding Author:

Name: Michelle McCabe

Address: Division of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, National Health Laboratory Services, Braamfontein, Johannesburg, 2000, South Africa.

Telephone: +27 11 4899160

Email: michelle.mccabe@nhls.ac.za

ORCID iD: 0000-0002-5066-1451

References

1. Iacopetta, B. Are there two sides to colorectal cancer? *Int. J. Cancer* **101**, 403–408 (2002).
2. Mik, M., Berut, M., Dziki, L., Trzcinski, R. & Dziki, A. Right-and left-sided colon cancer-clinical and pathological differences of the disease entity in one organ. *Arch. Med. Sci.* **13**, 157–162 (2017).
3. Lim, D. R., Kuk, J. K., Kim, T. & Shin, E. J. Comparison of oncological outcomes of right-sided colon cancer versus left-sided colon cancer after curative resection. *Med. (United States)* **96**, 1–7 (2017).
4. Benedix, F. *et al.* Comparison of 17,641 patients with right- and left-sided colon cancer: differences in epidemiology, perioperative course, histology, and survival. *Dis. Colon Rectum* **53**, 57–64 (2010).

5. Hemminki, K. *et al.* Tumor location and patient characteristics of colon and rectal adenocarcinomas in relation to survival and TNM classes. *BMC Cancer* **10**, 2–7 (2010).
6. Holch, J. W., Ricard, I., Stintzing, S., Modest, D. P. & Heinemann, V. The relevance of primary tumour location in patients with metastatic colorectal cancer: A meta-analysis of first-line clinical trials. *Eur. J. Cancer* **70**, 87–98 (2017).
7. Baran, B. *et al.* Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. *Gastroenterol. Res.* **11**, 264–273 (2018).
8. Medhanie, G. A. *et al.* Cancer incidence profile in sub-Saharan African-born blacks in the United States: Similarities and differences with US-born non-Hispanic blacks. *Cancer* **123**, 3116–3124 (2017).
9. Walker, A. R. P. & Segal, I. Colorectal cancer in an African city population in transition. *Eur. J. Cancer Prev.* **11**, 187–191 (2002).
10. Katsidzira, L. *et al.* The shifting epidemiology of colorectal cancer in sub-Saharan Africa. *Lancet Gastroenterol. Hepatol.* **2**, 377–383 (2017).
11. Center, M. M., Jemal, A., Smith, R. A. & Ward, E. Worldwide variations in colorectal cancer. *CA Cancer J Clin* **59**, 366–78 (2009).
12. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356 (2015).
13. Kim, S. Y. & Kim, T. Il. Serrated neoplasia pathway as an alternative route of colorectal cancer carcinogenesis. *Intest. Res.* **16**, 358–365 (2018).
14. Stintzing, S., Tejpar, S., Gibbs, P., Thiebach, L. & Lenz, H. J. Understanding the role of primary tumour localisation in colorectal cancer treatment and outcomes. *Eur. J. Cancer* **84**, 69–80 (2017).
15. Samowitz, W. S. *et al.* Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology* **129**, 837–845 (2005).
16. Muller, M., Ibrahim, A. & Arends, M. Molecular pathological classification of colorectal cancer. *Virchows Arch.* **469**, 125–134 (2016).
17. Ogino, S. *et al.* CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut* **58**, 90–96 (2009).
18. Hile, S. E., Shabashev, S. & Eckert, K. A. Tumor-specific microsatellite instability: Do distinct mechanisms underlie the MSI-L and EMAST phenotypes? *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **743–744**, 67–77 (2013).
19. Mori, Y. *et al.* The impact of microsatellite instability on the molecular phenotype of colorectal tumors. *Cancer Res.* **63**, 4577–4582 (2003).
20. Pawlik, T. M., Raut, C. P. & Rodriguez-bigas, M. A. Colorectal carcinogenesis : MSI-H versus MSI-L. **20**, 199–206 (2004).
21. Torshizi Esfahani, A., Seyedna, S. Y., Nazemalhosseini Mojarad, E., Majd, A. & Asadzadeh Aghdaei, H. MSI-L/EMAST is a predictive biomarker for metastasis in colorectal cancer patients. *J. Cell. Physiol.* **234**, 13128–13136 (2018).
22. Lee, S. Y. *et al.* Low-level microsatellite instability as a potential prognostic factor in sporadic colorectal cancer. *Med. (United States)* **94**, 1–8 (2015).
23. Mojarad, E. N. *et al.* Low Level of Microsatellite Instability Correlates with Poor Clinical Prognosis in Stage II Colorectal Cancer Patients. *J. Oncol.* **2016**, 1–9 (2016).
24. Wright, C. M. *et al.* Low level microsatellite instability may be associated with reduced cancer specific survival in sporadic stage C colorectal carcinoma. *Gut* **54**, 103–108 (2005).

25. Jass, J. R. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology* **50**, 113–130 (2007).
26. Balaguer, F. *et al.* Colorectal cancers with microsatellite instability display unique miRNA profiles. *Clin. Cancer Res.* **17**, 6239–6249 (2011).
27. Hrašovec, S. & Glavač, D. MicroRNAs as novel biomarkers in colorectal cancer. *Front. Genet.* **3**, 1–9 (2012).
28. Margetis, N., Kotsinas2, A. & T., M.-S. The Promising Role of Let-7 Microrna in Colorectal Cancer: Practical Points for Clinicians. *Clin. Oncol.* **3**, 8 (2018).
29. Cui, E. H. *et al.* Serum microRNA 125b as a diagnostic or prognostic biomarker for advanced NSCLC patients receiving cisplatin-based chemotherapy. *Acta Pharmacol. Sin.* **34**, 309–313 (2013).
30. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356 (2015).
31. Cronjé, L., Paterson, A. C. & Becker, P. J. Colorectal cancer in South Africa: A heritable cause suspected in many young black patients. *South African Med. J.* **99**, 103–106 (2009).
32. Prodehl, L., Bebington, B., Fabian, J., Singh, E. & Ruff, P. COLORECTAL CANCER IN A SOUTH AFRICA URBAN SETTING - A PRELIMINARY ANALYSIS. *S Afr J Surg* **55**, 58 (2017).
33. Bebington, B. *et al.* Design and methodology of a study on colorectal cancer in Johannesburg, South Africa. *JGH Open* **2**, 139–143 (2018).
34. McCabe, M., Perner, Y., Magobo, R., Mirza, S. & Penny, C. Descriptive epidemiological study of South African colorectal cancer patients at a Johannesburg Hospital Academic institution. *JGH Open* **4**, 360–367 (2020).
35. McCabe, M. *et al.* Microsatellite Instability assessment in Black South African Colorectal Cancer patients reveal an increased incidence of suspected Lynch syndrome. *Sci. Rep.* **9**, 1–10 (2019).
36. Haghighi, M. M. *et al.* Frequent MSI mononucleotide markers for diagnosis of hereditary nonpolyposis colorectal cancer. *Asian Pac. J. Cancer Prev.* **11**, 1033–1035 (2010).
37. Pyatt, R. *et al.* Implications for Microsatellite Instability Testing. *Cancer Res.* **155**, 349–353 (1999).
38. Suraweera, N. *et al.* Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* **123**, 1804–1811 (2002).
39. Buhard, O. *et al.* Multipopulation Analysis of Polymorphisms in Five Mononucleotide Repeats Used to Determine the Microsatellite Instability Status of Human Tumors. *J. Clin. Oncol.* **24**, 241–251 (2005).
40. Brennetot, C. *et al.* Mononucleotide repeats BAT-26 and BAT-25 accurately detect MSI-H tumors and predict tumor content: Implications for population screening. *Int. J. Cancer* **113**, 446–450 (2005).
41. Farraj, F. A., Sabbagh, H., Aridi, T., Fakhruddin, N. & Farhat, F. Signet Ring Cell Carcinoma of the Colon in Young Adults: A Case Report and Literature Review. *Case Rep. Oncol. Med.* **2019**, 1–8 (2019).
42. Wei, Q. *et al.* Clinicopathologic and molecular features of colorectal adenocarcinoma with signet-ring cell component. *PLoS One* **11**, 1–12 (2016).
43. Hamilton, S. *et al.* Tumours of the colon and rectum: Chapter 8. in *WHO Classification of Tumours of the Digestive System 2010* (eds. Bosman, F. T., Carneiro, F., Hruban, R. H. & D, T. N.) 137–138 (IARC: Lyon 2010, 2010).
44. Hartman, D. J. *et al.* Signet Ring Cell Colorectal Carcinoma. *Am. J. Surg. Pathol.* **37**, 969–977 (2013).
45. Schneider, N. I. & Langner, C. Prognostic stratification of colorectal cancer patients: Current perspectives. *Cancer Manag. Res.* **6**, 291–300 (2014).
46. Fleming, M., Ravula, S., Tatishchev, S. F. & Wang, H. L. Colorectal carcinoma: Pathologic aspects. *J. Gastrointest. Oncol.* **3**, 153–173 (2012).

47. Devaraj, B. *et al.* Relationship of EMAST and Microsatellite Instability Among Patients with Rectal Cancer. *J. Gastrointest. Surg.* **14**, 1521–1528 (2010).
48. Lee, S. Y. *et al.* Microsatellite instability, EMAST, and morphology associations with T cell infiltration in colorectal neoplasia. *Dig. Dis. Sci.* **57**, 72–78 (2012).
49. Garcia, M. *et al.* Association between recurrent metastasis from stage II and III primary colorectal tumors and moderate microsatellite instability. *Gastroenterology* **143**, 48–50 (2012).
50. Kohonen-Corish, M. R. J. *et al.* Low microsatellite instability is associated with poor prognosis in stage C colon cancer. *J. Clin. Oncol.* **23**, 2318–2324 (2005).
51. Buckowitz, A. *et al.* Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. *Br. J. Cancer* **92**, 1746–1753 (2005).
52. Rozek, L. S. *et al.* Tumor-Infiltrating lymphocytes, Crohn's-like lymphoid reaction, and survival from colorectal cancer. *J. Natl. Cancer Inst.* **108**, 1–8 (2016).
53. Posch, F. *et al.* Maturation of tertiary lymphoid structures and recurrence of stage II and III colorectal cancer. *Oncoimmunology* **7**, (2018).
54. Pastrello, C. *et al.* Stability of BAT26 in tumours of hereditary nonpolyposis colorectal cancer patients with MSH2 intragenic deletion. *Eur. J. Hum. Genet.* **14**, 63–68 (2006).
55. Jaskowski, L. *et al.* Stability of BAT26 in Lynch syndrome colorectal tumours [1]. *Eur. J. Hum. Genet.* **15**, 139–141 (2007).
56. Arnold, S. *et al.* Classifying MLH1 and MSH2 variants using bioinformatic prediction, splicing assays, segregation and tumor characteristics Sven. *Hum Mutat.* **30**, 757–770 (2009).
57. Thompson, B. A., Martins, A. & Spurdle, A. B. A review of mismatch repair gene transcripts: Issues for interpretation of mRNA splicing assays. *Clin. Genet.* **87**, 100–108 (2015).
58. Carethers, J. M. *et al.* Influence of race on microsatellite instability and CD8+ T cell infiltration in colon cancer. *PLoS One* **9**, 1–7 (2014).
59. Wang, J., Liu, J., Tian, F., Zhan, Y. & Kong, D. Cyclin-dependent kinase 9 expression and its association with CD8+ T cell infiltration in microsatellite-stable colorectal cancer. *Oncol. Lett.* **18**, 6046–6056 (2019).
60. Frumento, G. *et al.* CD117 (c-Kit) is expressed during CD8+ T cell priming and stratifies sensitivity to apoptosis according to strength of TCR engagement. *Front. Immunol.* **10**, (2019).
61. Johnson, S. M. *et al.* RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635–647 (2005).
62. Madison, B. B. *et al.* Let-7 Represses Carcinogenesis and a Stem Cell Phenotype in the Intestine via Regulation of Hmga2. *PLoS Genet.* **11**, 1–21 (2015).
63. Yong, S. L. & Dutta, A. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes Dev.* **21**, 1025–1030 (2007).
64. Liu, T. P. *et al.* Down-regulation of let-7a-5p predicts lymph node metastasis and prognosis in colorectal cancer: Implications for chemotherapy. *Surg. Oncol.* **25**, 429–434 (2016).
65. Chirshev, E., Oberg, K. C., Ioffe, Y. J. & Unternaehrer, J. J. Let-7 as biomarker, prognostic indicator, and therapy for precision medicine in cancer. *Clin. Transl. Med.* **8**, (2019).
66. Gao, S., Zhao, Z. Y., Wu, R., Zhang, Y. & Zhang, Z. Y. Prognostic value of microRNAs in colorectal cancer: A meta-analysis. *Cancer Manag. Res.* **10**, 907–929 (2018).
67. Mizuno, R., Kawada, K. & Sakai, Y. The Molecular Basis and Therapeutic Potential of Let-7 MicroRNAs against Colorectal Cancer. *Can. J. Gastroenterol. Hepatol.* **2018**, (2018).

68. King, C. E. *et al.* LIN28B fosters colon cancer migration, invasion and transformation through let-7-dependent and-independent mechanisms. *Oncogene* **30**, 4185–4193 (2011).
69. Volinia, S. *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2257–2261 (2006).
70. Fang, W. *et al.* Detection of let-7a microRNA by real-time PCR in colorectal cancer: A single-centre experience from China. *J. Int. Med. Res.* **35**, 716–723 (2007).
71. Michael, M. Z., O'Connor, S. M., Van Holst Pellekaan, N. G., Young, G. P. & James, R. J. Reduced Accumulation of Specific MicroRNAs in Colorectal Neoplasia. *Mol. Cancer Res.* **1**, 882–891 (2003).
72. Kok, M. G. M. *et al.* Small sample sizes in high-throughput miRNA screens: A common pitfall for the identification of miRNA biomarkers. *Biomol. Detect. Quantif.* **15**, 1–5 (2018).
73. Kets, C. M. *et al.* Unfavorable pathological characteristics in familial colorectal cancer with low-level microsatellite instability. *Mod. Pathol.* **19**, 1624–1630 (2006).
74. Azzoni, C. *et al.* Sporadic colorectal carcinomas with low-level microsatellite instability: A distinct subgroup with specific clinicopathological and molecular features. *Int. J. Colorectal Dis.* **26**, 445–453 (2011).
75. Oliart S, Martínez-Santos C, Moreno-Azcoita M, Cerquella C, Nejda N, Daimiel L, Iglesias D, Fernández-Peralta AM, G.-A. J. Do MSI-L sporadic colorectal tumors develop through " mild mutator pathway "? *Am. J. Clin. Oncol.* **29**, 364–70 (2006).
76. Ozaki, K. *et al.* Heterogeneous microsatellite instability observed within epithelium of ulcerative colitis. *Int. J. Cancer* **119**, 2513–2519 (2006).
77. Ashktorab, H. *et al.* Identification of novel mutations by exome sequencing in African American colorectal cancer patients. *Cancer* **121**, 34–42 (2015).
78. Carethers, J. M., Koi, M. & Tseng-Rogenski, S. S. EMAST is a form of microsatellite instability that is initiated by inflammation and modulates colorectal cancer progression. *Genes (Basel).* **6**, 185–205 (2015).
79. Grady, W. M. & Carethers, J. M. Genomic and Epigenetic Instability in Colorectal Cancer Pathogenesis. *Cancer* **135**, 1079–1099 (2008).
80. Koi, M., Tseng-Rogenski, S. S. & Carethers, J. M. Inflammation-associated microsatellite alterations: Mechanisms and significance in the prognosis of patients with colorectal cancer. *World J. Gastrointest. Oncol.* **10**, 1–14 (2018).
81. Cole, B. F. *et al.* Aspirin for the chemoprevention of colorectal adenomas: Meta-analysis of the randomized trials. *J. Natl. Cancer Inst.* **101**, 256–266 (2009).
82. Cook, N. R., Lee, I. M., Zhang, S. M., Moorthy, M. V. & Buring, J. E. Alternate-day, low-dose aspirin and cancer risk: Long-term observational follow-up of a randomized trial. *Ann. Intern. Med.* **159**, 77–85 (2013).
83. Arber, N. *et al.* Celecoxib for the prevention of colorectal adenomatous polyps. *N. Engl. J. Med.* **355**, 885–895 (2006).
84. Baron, J. A. *et al.* A Randomized Trial of Rofecoxib for the Chemoprevention of Colorectal Adenomas. *Gastroenterology* **131**, 1674–1682 (2006).
85. Yamauchi, T. *et al.* The potential for a selective cyclooxygenase-2 inhibitor in the prevention of liver metastasis in human colorectal cancer. *Anticancer Res.* **23**, 245–249 (2003).
86. Chan, A. T., Ogino, S. & Fuchs, C. S. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. *Sci. Transl. Med.* **356**, 2131–42 (2007).
87. Hamada, T. *et al.* Aspirin use and colorectal cancer survival according to tumor CD274 (programmed cell death 1 ligand 1) expression status. *J. Clin. Oncol.* **35**, 1836–1844 (2017).

Tables

Table 1: Descriptive analysis of MSS CRC cases diagnosed at CMJAH between (2011-2015): Categorized by site: LCC *vs* RCC and Ethnicity: Indigenous African (IA) *vs* Other Ethnic Group (OEG)

	Total number cases(%)	Left-sided Colon Cancer (LCC)		Statistical analysis:	Right-sided Colon Cancer (RCC)		Statistical analysis:
		IA	OEG		IA	OEG	
Frequency	369	154 (42)	108(29)		63(17)	44(12)	
Demographical data							
GENDER	369	154	108	P=0.0111*	63	44	P=0.6932
Male	209(57)	98(64)	51(47)		34(54)	26(59)	
Female	160(43)	56(36)	57(53)		29(46)	18(41)	
AGE	365	152	107	P<0.0001***	62	44	P=0.0103*
Min-Max	15-92	20-90	28-92		15-79	25-86	
Mean ± SD	57±14	53±15	62 ±13		54±13	60 ±11	
Median	59	54***	62***		55*	61*	
P25-P75 (Interquartile Range)	47-67	41-65	55-72		47-64	54-68	
95% CI	[55-58]	[51-56]	[60-65]		[51-57]	[57-65]	
Categorical Age	365	152	107	P<0.0001***	62	44	P<0.0001***
≤50years	107 (29)	62(41)	15(14)		24(39)	6(14)	
>50 years	258(71)	90(59)	92(86)		38(61)	38(86)	
Histological characteristics							
TUMOUR SUBTYPE	360	152	104	P=0.0257*	61	43	P=0.2608
Invasive Adenocarcinoma	329(91)	138(91)	100(96)		54(89)	35(81)	
Mucinous Adenocarcinoma	17(5)	4(3)	4(4)		3(5)	6(14)	
Signet Ring Cell Adenocarcinoma	14(4)	10(6)	0(0)		4(6)	2(5)	
TUMOUR SITE:	359	151	105	P = 0.0221*	61	42	P = 0.2431
Left / Right							
Splenic Flexure / Hepatic Flexure	14(4)	5(3)	0(0)		6(10)	3(7)	
Descending colon / Ascending Colon	58(16)	21(14)	11(10)		11(18)	15(36)	
Sigmoid / Transverse Colon	83(23)	46(31)	21(20)		10(16)	6(14)	
Rectum / Caecum	204(57)	79(52)	73(70)		34(56)	18(43)	
TUMOUR GRADE	348	148	95	P = 0.0930	61	44	P = 1.0000
Low Grade (LG)	316(91)	132(89)	91(96)		54(89)	39(89)	

High Grade (HG)	32(9)	16(11)	4(4)		7(11)	5(11)	
AJCC TNM STAGING	240	85	61	P=0.0922	57	37	P= 0.4995
I-II	92(38)	31(38)	31(51)		20(36)	10(27)	
III-IV	148(62)	54(62)	30(49)		37(64)	27(73)	
TUMOUR INFILTRATING LYMPHOCYTES (TIL)	230	81	60	P = 0.8512	54	35	P=0.6591
None	156(68)	58(72)	44(73)		34(63)	20(57)	
Mild-moderate	74(32)	23(28)	16(27)		20(37)	15(43)	
CROHN'S LIKE INFLAMMATORY RESPONSE	230	81	60	P = 0.4406	54	35	P=0.3392
None	168(73)	62(77)	42(70)		41(76)	23(66)	
Mild-moderate	62(27)	19(23)	18(30)		13(24)	12(34)	
LYMPHATIC INVASION	285	115	76	P = 0.2921	56	38	P=0.2896
Absent	203(71)	85(74)	62(82)		36(64)	20(53)	
Present	82(29)	30(26)	14(18)		20(36)	18(47)	
VENOUS INVASION	232	82	59	P = 0.3937	55	36	P = 1.0000
Absent	181(78)	68(83)	45(76)		41(75)	27(75)	
Present	51(22)	14(17)	14(24)		14(25)	9(25)	
PERINEURAL INVASION	241	91	60	P = 0.1110	54	36	P = 1.0000
Absent	187(78)	67(74)	51(85)		41(76)	28(78)	
Present	54(22)	24(26)	9(15)		13(24)	8(22)	
POLYPS	80	31	19	P = 0.5516	17	14	P = 0.7131
Tubular Adenoma (TA)	49(58)	21(66)	11(52)		10(59)	7(50)	
Tubulovillous Adenoma (TVA)	31(37)	10(31)	8(38)		6(35)	7(50)	
MSS/MSI-L PCR CONFIRMED CASES	244	105	85	P=0.0442*	28	26	P=0.7471
MSS	202(83)	83(79)	77(91)		21(75)	21(81)	
MSI-L	42(17)	22(21)	8(9)		7(25)	5(19)	
MSI PCR MARKERS (MSI-L)	42	22	8	P=0.0026**	7	5	P=0.2222
1 single unstable marker	20(48)	11(50)	4(50)	BAT25/26 *Vs* NR21/24/27	4((57)	1(20)	BAT25/26 *Vs* NR21/24/27
BAT25	14(33)	11(50)	0(0)		2(29)	1(20)	
BAT26	5(12)	0(0)	3(38)		0(0)	2(40)	
NR21							

NR24	3(7)	0(0)	1(12)	1(14)	1(20)
NR27	0(0)	0(0)	0(0)	0(0)	0(0)

Table 2: Descriptive analysis of MSS and MSI-L CRC cases diagnosed at CMJAH (2011-2015): Categorized by MSI status: MSS *vs* MSI-L CRC and site: LCC *vs* RCC.

Figures

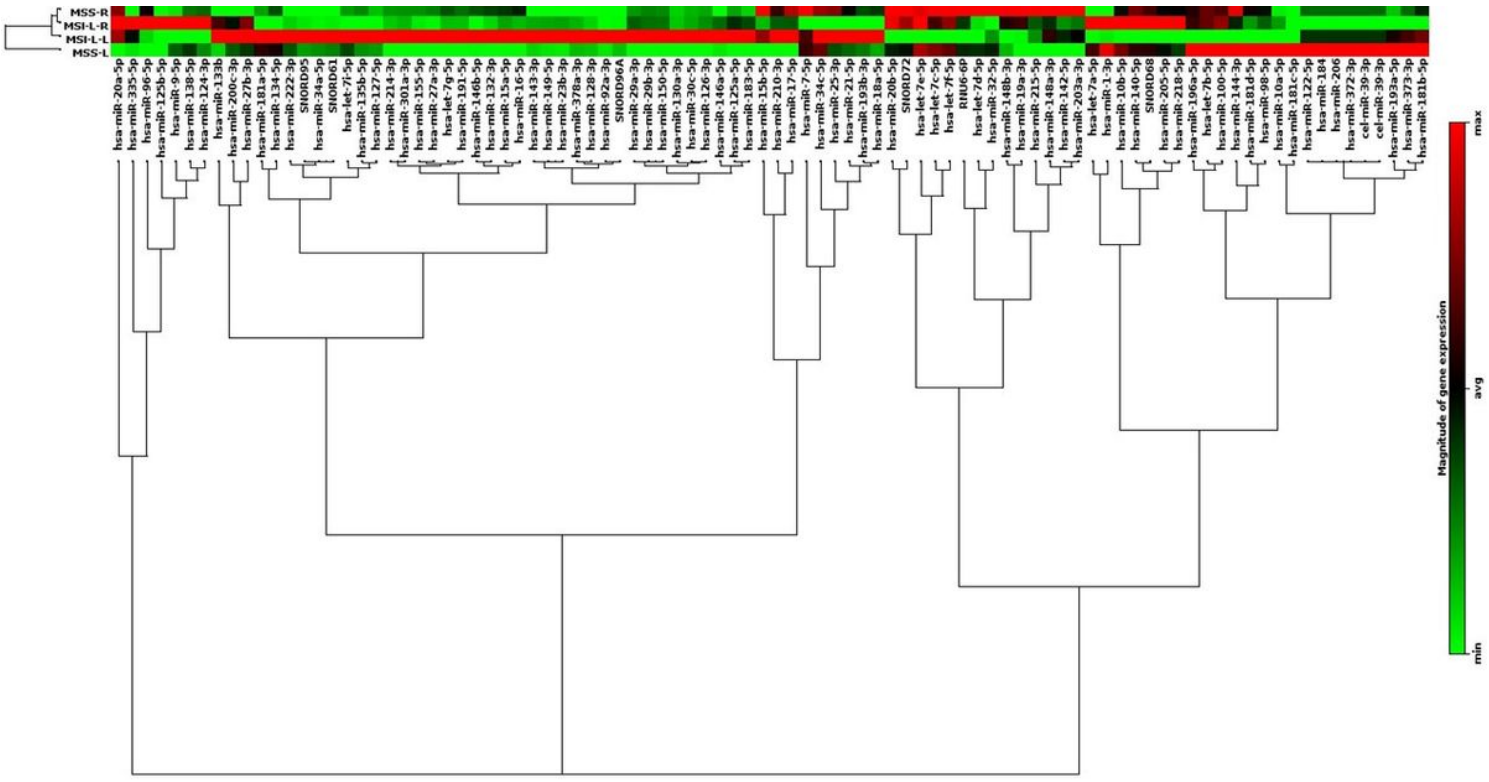

Figure 1

Cluster gram illustrating miRNA expression profiles of subgroups MSI-L RCC, MSI-L LCC, and MSS-RCC compared to subgroup MSS-LCC: The cluster gram depicts different molecular subgroups based on miRNA expression patterns. The red colour presents upregulated expression, green is downregulated, and brown to black, average expression. Let-7a-5p was downregulated for MSI-L LCC and MSS RCC when compared to MSS LCC. Isoforms let-7e-5p, let-7c-5p and let-7f-5p were also downregulated in MSI-L LCC when compared to remaining 3 subgroups. MiR-125b was downregulated in MSI-L LCC, MSS RCC, and MSS LCC, however upregulated in the MSI-L RCC subgroup. MiR-125a was considerably upregulated for MSI-L LCC compared to MSS LCC by approximately 4.3K log fold. Overall this data demonstrates different expression patterns, suggesting 4 distinct molecular subgroups.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinformation.docx

	Number of cases (%)	MSI-L CRC		MSS CRC		
	CRC 2011-2015	A: LCC	B: RCC	C: LCC	D: RCC	Statistical analysis:
Frequency/ Prevalence	244	30(13)	12(6)	160(64)	42(17)	
GENDER	244	30	12	160	42	A vs B: P=1.0000
Male	134(55)	16(53)	6(50)	88(55)	24(57)	A vs C: P=0.8414
Female	110(45)	14(47)	6(50)	72(45)	18(43)	A vs D: P=0.6299
						B vs C: P=1.0000
						B vs D: P=0.7573
						C vs D: P=0.7227
AGE	244	29	12	160	42	A vs B: P=0.5250
Min-Max	15-92	24-91	37-84	20-92	15-80	A vs C: P=0.6944
Mean ± SD	57±14	57±17	59 ±13	58±15	57 ±13	A vs D: P=0.7750
Median	59	57	59	59	58	B vs C: P=0.7068
P25-P75 (IQR)	48-69	40-72	60	48-69	53-66	B vs D: P=0.7002
95% CI	[56-59]	[50-63]	52-69 [52-67]	[55-60]	[53-62]	C vs D: P=0.9170 (q value=0.8238) Q=0.05
ETHNIC GROUPS	244	30	12	160	42	A vs B: P=0.4635
Indigenous African	133(55)	22(76)	7(58)	83(47)	21(50)	A vs C: P=0.0442*
Other Ethnic groups	111(45)	8(24)	5(42)	77(53)	21(50)	A vs D: P=0.0550
						B vs C: P=0.7693
						B vs D: P=0.7471
						C vs D: P=0.8636
AJCC TNM STAGING	130	9	8	76	37	A vs B:

I-II	55(42)	1(11)	4(50)	38(50)	12(32)	P=0.1312
III-IV	75(58)	8(89)	4(50)	38(50)	25(68)	A vs C: P=0.0348* A vs D: P=0.4100 B vs C: P=1.0000 B vs D: P=0.4270 C vs D: P=0.1063
CROHN'S LIKE INFLAMMATORY RESPONSE	125	9	8	73	35	A vs B: P=0.0498*
Mild-moderate	36((29)	6(67)	1(12)	19(26)	10(29)	A vs C: P=0.0205*
Absent	89(71)	3(33)	7(88)	54(74)	25(71)	A vs D: P=0.0534 B vs C: P=0.6720 B vs D: P=0.6563 C vs D: P=0.8188