


ARTICLE



Association of microsatellite instability (MSI) status with the 5-year outcome and genetic ancestry in a large Brazilian cohort of colorectal cancer

Gustavo Noriz Berardinelli^{1,2}, Ronílson Durães^{1,3}, Allini Mafra da Costa⁴, Arinilda Bragagnoli³, Marco Antônio de Oliveira⁵, Rui Pereira^{6,7}, Cristovam Scapulatempo-Neto¹, Denise Peixoto Guimarães^{1,8} and Rui Manuel Reis^{1,2,9,10} 

© The Author(s), under exclusive licence to European Society of Human Genetics 2022

Colorectal cancer (CRC) has a high incidence and mortality worldwide. Microsatellite instability (MSI) is crucial in CRC, with distinct molecular and clinicopathological features in patients. Nowadays, it is a predictive marker for immunotherapy. We proposed to evaluate the 5-year outcome of MSI status in 1002 Brazilian CRC, and associate it with genetic ancestry, molecular and clinicopathological features. MSI evaluation was performed using molecular markers. MSI+ tumors were analyzed for alterations in 23 MSI-targeted genes. Genetic ancestry was evaluated using an Ancestry-Informative markers panel. MSI status was analyzed in relation to CRC specific survival and other clinical and genetic variables. MSI+ status was observed in 10.5% of cases. MSI+ status was significantly associated with the anatomic site right colon, mucinous histological type, clinical stage II, histological grade III/undifferentiated, no recurrence of disease, and live cases without cancer. No association of MSI status with genetic ancestry components was observed. MSI-targeted genes analyses showed the most frequently altered genes: *ATM*, *EGFR*, *MRE11*, *ROCK1*, and *TGFBR1*. There was a statistically significant difference in cancer-specific survival between cases according to MSI status. This study constitutes the most comprehensive analyses of the MSI impact on the Brazilian CRC. MSI+ frequency in Brazilian CRC agreed with the literature and was associated with several clinicopathological features related with less aggressive tumors, independently of their genetic ancestry.

European Journal of Human Genetics (2022) 30:824–832; <https://doi.org/10.1038/s41431-022-01104-y>

INTRODUCTION

Colorectal cancer (CRC) is a common disease worldwide, tending to rise uniformly with increasing human development index in many countries [1, 2]. In Brazil, it is the second most common cause of cancer for men and women [3]. The rise of CRC incidence rates observed in the last decade is due to population aging, increasing smoking rates, poor dietary habits, low physical activity, and the absence of widespread screening programs [4, 5]. The Brazilian CRC mortality rate is increasing for both sexes when comparing data across the last decades and data from Latin American countries [6].

CRC is a heterogeneous molecular disease, which leads to distinct clinicopathological features and patient outcome [4]. The cumulative acquisition of genetic alterations leads to a progressive tumorigenesis process [4]. These alterations are linked to three main molecular groups: chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylation phenotype (CIMP), and these genetic pathways are involved in the development of CRC affecting oncogenes, tumor suppressor genes, and DNA repair mechanisms [7, 8].

Nowadays, CRC has also been subtyped molecularly in four consensus molecular subtypes (CMS) based on gene expression with distinguishing features: CMS1 (microsatellite instability immune, 14%), hypermutated, microsatellite unstable and robust immune activation; CMS2 (canonical, 37%), epithelial, marked WNT and MYC signaling activation; CMS3 (metabolic, 13%), epithelial and evident metabolic dysregulation; and CMS4 (mesenchymal, 23%), prominent transforming growth factor-beta activation, stromal invasion and angiogenesis [9].

Microsatellite instability (MSI) is a crucial feature of a subset of CRC [4, 8] and is a molecular marker for defects in the mismatch repair system. It occurs when Mismatch repair (MMR) proteins (MLH1, MLH3, MSH2, MSH3, MSH6, PMS1, and PMS2) are absent due to mutations or promoter hypermethylation in hereditary and sporadic cancer, respectively [8, 10]. These proteins are essential to repair base-base mismatches occurring during DNA replication; thus, their loss guides DNA replication errors accumulation, mostly in microsatellites genomic areas [11]. As alterations occur in a random matter, findings suggest a different progression of each

¹Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, São Paulo, Brazil. ²Molecular Diagnostic Laboratory, Barretos Cancer Hospital, Barretos, São Paulo, Brazil.

³Department of Oncology, Barretos Cancer Hospital, Barretos, São Paulo, Brazil. ⁴Population-based Cancer Registry of Barretos Region, Barretos Cancer Hospital, Barretos, São Paulo, Brazil.

⁵Nucleus of Epidemiology and Biostatistics, Barretos Cancer Hospital, Barretos, São Paulo, Brazil. ⁶IPATIMUP – Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto, Portugal.

⁷IS - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal. ⁸Department of Endoscopy, Barretos Cancer Hospital, Barretos, São Paulo, Brazil.

⁹Life and Health Sciences Research Institute, University of Minho, Gualtar Campus, 4710-057 Braga, Portugal. ¹⁰ICVS/3B's-PT Government Associate Laboratory, Gualtar Campus, 4710-057 Braga/Guimarães, Portugal. [✉]email: ruireis.hcb@gmail.com

Received: 8 September 2021 Revised: 4 February 2022 Accepted: 12 April 2022

Published online: 26 April 2022

MSI-positive (MSI+) tumor, a model in which MSI mutator phenotype develops in gradual steps by successive alterations of different MSI-target genes [12, 13].

MSI+ CRCs are associated with clinical features, such as proximal location, poorly differentiated histology, intense lymphocytic infiltration, favorable prognosis, and lower risk of metastasis [4, 11–13]. Moreover, evidence suggests that MSI+ patients are less responsive to 5-fluorouracil-based chemotherapy (5-FU), and are more responsive to irinotecan-based regimens [4, 11, 14]. Importantly, MSI was the first agnostic biomarker approved by the FDA (Food and Drug Administration) to select patients to be treated with immunotherapy treatments [15].

It is reported that African Americans with CRC are typically diagnosed at a younger age than European Americans and display high mortality rates even at early stages of CRC [16, 17]. Likewise, our group recently assessed the genetic ancestry profile of 1000 patients and observed that patients with high African genetic ancestry proportions developed cancer at a younger age [18]. A recent meta-analysis of MSI frequency and ethnicity in CRC did not observe significant differences among the North American population [19].

Few studies have comprehensively described and characterized the main clinicopathological features of highly admixture populations, such as Brazilian CRC patients [18, 20], especially concerning MSI status and its clinical impact. Therefore, this study aimed to evaluate the long-term outcome of MSI status of 1002 Brazilian CRC and associate it with genetic ancestry, molecular and clinicopathological features.

METHODS

Participants

The present study included 1002 patients diagnosed between 2010 and 2014 with CRC at Barretos Cancer Hospital. The median follow-up of our cases was 62.0 months. During the inclusion period, patients diagnosed with Lynch Syndrome were excluded [21]. Clinicopathological and treatment data was recently reported [18]. AJCC (American Joint Committee on Cancer) cancer staging system, 8th Edition classification was applied. The study was evaluated and approved by Institutional Ethics Committee (protocol number: 600/2012 - CAAE: 02468812.30000.5437).

DNA isolation and microsatellite instability

Tumor DNA was isolated using QIAamp DNA Micro Kit (Qiagen, Hilden, Germany). MSI analysis of the current series was recently reported by our group [22]. Briefly, MSI evaluation was performed using a multiplex PCR comprising six quasi-monomorphic mononucleotide repeat markers (BAT25, BAT26, NR21, NR24, NR27, and HSP110). Cases with two or more markers out of the quasimonomorphic variation range (QMVR) were classified as MSI-positive (MSI+), and cases without markers out of QMVR were classified as MSI-negative (MSI–), as reported [23].

MSI-target genes

We analyzed 23 MSI-target genes that contained microsatellite regions in their constitution and were previously reported to be important in CRC carcinogenesis [12, 13]. Multiplex PCR was performed for evaluation in MSI+ cases in a 23-gene panel: *TCF4*, *XRCC2*, *MBD4*, *MRE11*, *ATR*, *MSH3*, *RAD50*, *MSH6*, *BAX*, *DNAPKc*, *BRCA1*, *BRAC2*, *WISP3*, *BLM*, *PTEN*, *ATM*, *TGFBR2*, *XPO5*, *TRBP2*, *EGFR*, *ABCC5*, *ROCK1*, and *GLYR1*, as previously described [24].

For multiplex PCR, 5 μ L of Qiagen multiplex PCR master mix (Qiagen, Germany), 1 μ L of primer mix (1 μ M), 3 μ L of water, and 1 μ L of DNA with a concentration of 50 ng/ μ L were used, with a 10 μ L final volume. Cycling was carried out under the following conditions: 95 °C for 15 min for initial DNA denaturation, followed by 30 cycles of 94 °C for 30 s, 55 °C for 90 s and 72 °C for 45 s that provided denaturation, annealing, and extension, respectively. The final extension was provided by the 72 °C stages for 40 min and a standby temperature of 4 °C.

PCR products were then prepared for capillary electrophoresis by adding 1 μ L of the amplified product, 8.7 μ L of Hi-Di Formamide (Applied Biosystems, Foster City, CA, USA), and 0.3 μ L of GeneScan 500 ROX size standard (Applied Biosystems, USA). The presence of alterations in

microsatellite regions of these genes was assessed by fragment analyzes in 3500 Genetic Analyzer (Applied Biosystems, USA). Data generated by 3500 Genetic Analyzer equipment were analyzed using GeneMapper software version 4.0 (Applied Biosystems, USA). No-altered status was given for tumor samples whose fragment size was similar to that of normal tissue. Altered status was given for tumor samples whose fragment size was different (expansion or contraction in microsatellite) than normal tissue.

Ancestry analysis

Ancestry determination of the present series was recently reported by our group [18]. Briefly, it was performed using 46-plex ancestry-informative markers (AIMs) among the most informative INDELs for four major population groups (African, European, Eastern Asian, and Amerindian) as previously described [25]. Ancestry proportions were then assessed using Structure v2.3.4 software [26, 27], considering each major population group as possible contributors to the current genetic makeup of Brazilians. Supervised analysis was performed to estimate ancestry membership ratios of individuals using HGDP-CEPH panel data as a reference for ancestral populations.

Statistical analysis

The sample was characterized using frequency and/or contingency tables for qualitative variables, and for quantitative variables were used measures of central tendency and dispersion (mean, median, standard deviation, minimum and maximum).

Chi-Square or Fisher's Exact Tests were used to verify MSI status association with demographic data, clinical-pathological characteristics, and genomic ancestry and the multiple comparisons between the columns were performed by Bonferroni method. Variables considered significant ($p < 0.20$) were selected to fit the Multiple Logistic Regression Model, through which we estimated the Odds Ratio by the final model that was composed of all variables that remained significant together with a level of significance of 5%.

Survival probability was estimated using the Kaplan-Meier method, and comparison between curves was performed using the Log-Rank test. Characteristics considered significant ($p < 0.20$) in that test were used to adjust Cox Proportional-Hazards Model by which we estimate Hazard Ratio (HR). To compose the final model, characteristics that remained significant ($p < 0.05$) were used together. Analyzes were performed using SPSS software version 27 (IBM Corp, Armonk, NY, USA).

RESULTS

Association analyses between MSI status and clinicopathological features

Herein, we analyzed the association of MSI status with patients' clinicopathological and ancestry features from 105 MSI+ and 897 MSI– CRC patients (Table 1). In a univariate analysis, we observed that MSI+ was significantly associated with tumors sited in the right colon, mucinous histological type, clinical stage II, histological grade III/Undifferentiated, no recurrence disease, and live cases without cancer (Table 1). No association of MSI status with genetic ancestry components (European, African, Eastern Asian, and Amerindian) nor with Brazilian origin of patients was observed (Table 1 and Fig. 1).

Next, a multivariate analysis was performed with all variables with a $p < 0.2$, obtained after univariate analysis with MSI status as the outcome. We observed that MSI+ was significantly associated with tumors in the right colon, histological grade III/undifferentiated, and clinical stage IV (Supplementary Table I).

MSI-target genes

In MSI+ cases, we further evaluate the status of a panel of 23 MSI-target genes (Fig. 2 and Supplementary Fig. 1). These MSI-target genes can be stratified by their function in DNA repair (*XRCC2*, *MBD4*, *MRE11*, *MSH3*, *MSH6*, *RAD50*, *DNAPKc*, *BRCA1*, *BRCA2*, *BLM*, *ATM*, and *ATR*), apoptosis (*BAX*), cell signaling (*EGFR*, *PTEN*, *TCF4*, and *TGFBR1*), microRNA regulation (*TRBP2* and *XPO5*), oxidation (*GLYR1*), adhesion/cytoskeleton (*WISP3* and *ROCK1*) and transport (*XPO5* and *ABCC5*). The five most altered genes

Table 1. Description of clinicopathological features for MSI status and association analyses between MSI status and clinicopathological features for CRC.

Variable	Category	MSI status		p value	Odds Ratio (CI)
		MSI- % (n)	MSI+ % (n)		
Sex	Male	52.5 (471)	46.7 (49)	0.30	1
	Female	47.5 (426)	53.3 (56)		1.3 (0.8–1.9)
Age at diagnose (in years)	<50	29.1 (261)a	27.6 (29)a	0.07	1
	≥50 e < 75	61.1 (548)a	55.2 (58)a		0.9 (0.6–1.5)
	≥75	9.8 (88)a	17.1 (18)b		1.8 (1.0–3.5)
Primary tumor site	Right colon	19.7 (177)a	69.5 (73)b	<0.001	1
	Left colon	49.4 (443)a	22.9 (24)b		0.1 (0.1–0.2)
	Rectum	30.9 (277)a	7.6 (8)b		0.1 (0.03–0.10)
Histological type	Adenocarcinoma	94.3 (846)a	86.7 (91)b	0.005	1
	Mucinous	5.2 (47)a	10.5 (11)b		2.18 (1.1–4.3)
	Others	0.5 (4)a	2.8 (3)b		6.97 (1.5–31.6)
Clinical stage (AJCC)	0/I	12.8 (115)a	9.5 (10)a	0.004	1
	II	36.1 (323)a	50.5 (53)b		1.9 (0.9–3.8)
	III	33.2 (297)a	33.3 (35)a		1.3 (0.6–2.8)
	IV	17.9 (160)a	6.7 (7)b		0.5 (0.2–1.4)
Histological grade	I/II	95.5 (845)a	77.5 (79)b	<0.001	1
	III/Undifferentiated	4.5 (40)a	22.5 (23)b		6.1 (3.5–10.8)
Angiolymphatic invasion	No	67.5 (553)	66.7 (68)	0.91	1
	Yes	32.5 (266)	33.3 (34)		1.0 (0.7–1.6)
Perineural invasion	No	85.4 (647)	87.5 (84)	0.65	1
	Yes	14.6 (111)	12.5 (12)		0.8 (0.4–1.6)
Presence of synchronous tumors	No	94.5 (848)	91.4 (96)	0.27	1
	Yes	5.5 (49)	8.6 (9)		1.6 (0.8–3.4)
Presence of recurrence	No	68.6 (615)	79.0 (83)	0.03	1
	Yes	31.4 (282)	21.0 (22)		0.58 (0.3–0.9)
Vital status	Alive without cancer	44.3 (397)a	61.9 (65)b	<0.001	1
	Alive with cancer	4.3 (39)a	8.6 (9)a		1.41 (0.6–3.0)
	Death by cancer	44.6 (400)a	21.0 (22)a		0.34 (0.2–0.6)
	Death by others causes	6.5 (58)a	8.6 (9)a		0.95 (0.4–2.0)
	Loss of follow up	0.3 (3)a	0 (0)a		0.87 (0.04–16.97)
African ancestry component	Low	35.3 (317)	28.6 (30)	0.40	1
	Intermediate	27.3 (245)	30.5 (32)		1.4 (0.8–2.3)
	High	37.3 (335)	41.0 (43)		1.3 (0.8–2.2)
European ancestry component	Low	31.4 (282)	32.4 (34)	0.70	1
	Intermediate	29.4 (264)	25.7 (27)		0.8 (0.5–1.4)
	High	39.1 (351)	41.9 (44)		1.0 (0.6–1.7)
Eastern Asian ancestry component	Low	44.4 (398)	44.8 (47)	0.20	1
	Intermediate	19.7 (177)	26.7 (28)		1.3 (0.8–2.2)
	High	35.9 (322)	28.6 (30)		0.8 (0.5–1.3)
Amerindian ancestry component	Low	40.0 (359)	42.9 (45)	0.30	1
	Intermediate	22.6 (203)	26.7 (28)		1.1 (0.7–1.8)
	High	37.3 (335)	30.5 (32)		0.8 (0.5–1.2)
Brazilian region of origin	South/Southeast	82.1 (736)	81.9 (86)	0.30	1
	Midwest	9.0 (81)	5.7 (6)		0.63 (0.27–1.50)
	North/Northeast	8.9 (80)	12.4 (13)		1.40 (0.74–2.60)

MSI-: cases without microsatellite instability. MSI+: cases with microsatellite instability. n: total number of cases. a, b: different letters show a statistically significant difference after multiple comparisons with the significance level correction by Bonferroni method. Significant associations are indicated in bold. CI confidential interval.

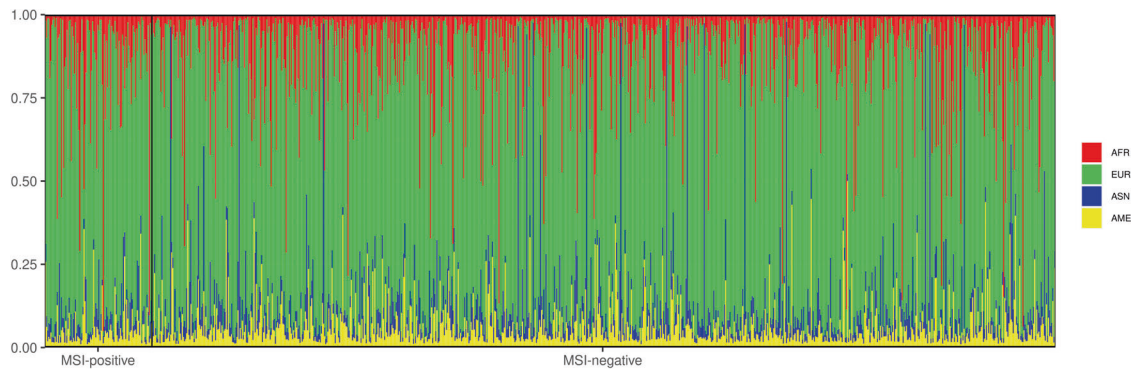


Fig. 1 Genetic ancestry analysis. Graphical representation of the genetic ancestry component of each case separated by MSI status (MSI-positive and MSI-negative). The ancestry analysis was performed using a set of 46 AIMs among the most informative INDELS for each population group and using the genetic data from the HGDP-CEPH panel as a reference. A supervised analysis was performed to estimate ancestry proportions of the individuals. Structure software runs considering $K = 4$ consisted of 100 000 burning steps followed by 100 000 Markov Chain Monte Carlo iterations. The option 'Use population Information to test for migrants' was used with the Admixture model, considering allele frequencies correlated, and updating allele frequencies using only individuals with POPFLAG = 1. The proportion in percentage of each ancestry component for each patient (columns) is represented by colors (Y axis): red – African, green – European, blue – Eastern Asian and yellow – Amerindian.

were: *ATM*, *EGFR*, *MRE11*, *ROCK1*, and *TGFBR11* (Fig. 2), corresponding gene functions related to DNA repair, adhesion/cytoskeleton, and cell signaling. We also found the absence of alterations in *BRCA1*, *BRCA2*, *XPO5*, and *XRCC2* (Fig. 2).

Impact of MSI status in patient survival

The 5-years cancer-specific survival (CSS) analyses of CRC cases were stratified by MSI status. MSI+ patients had significantly higher 5-years probability of survival than MSI- cases, 77.7%, and 60.5%, respectively (Supplementary Table II and Supplementary Fig. 2).

Several significant associations were observed between CSS and MSI + CRC patients' features, including better survival probability for patients with clinical stage II and III, absence of angiolymphatic and perineural invasions and no recurrence of disease (Table 2). For MSI- cases, better survival probability was observed, including female patients, clinical stage 0/I and II, tumors with mucinous histological type and I/II histological grade, absence of angiolymphatic and perineural invasions, no recurrence of disease, and patients underwent treatments such as neoadjuvant and adjuvant chemotherapy and radiotherapy (Table 2).

Multivariate analysis for CSS showed that in MSI+ cases, tumors with clinical stage IV, presence of perineural invasion, and recurrence of disease were associated with an increased relative risk of death by cancer ($p < 0.05$). Additionally, patients with age of diagnoses in years between ≥ 50 to < 75 were associated with a lower risk of death (Table 3). For MSI- cases, the analysis showed that: tumors with clinical stage II, III, and IV, presence of angiolymphatic invasion, and recurrence of disease were associated with an increased relative risk of death by cancer ($p < 0.05$), whereas female gender cases and the adjuvant chemotherapy treatment were associated with a lower risk of death (Table 3).

Given the importance that the clinical stage has as a significant prognostic factor and treatment determinant, estimates of CSS stratified by MSI status for all stages were performed. Overall, there was a statistically significant difference for clinical stage III, where MSI+ patients showed a better 5-years probability than MSI- patients (Table 4 and Supplementary Fig. 3). Regarding ancestry components, Kaplan–Meier estimates of CSS for patients by MSI status were performed, and no statistically significant differences were observed (Supplementary Table III).

Then, we analyzed the impact of the MSI-target genes on patient survival. For these analyses, altered cases were considered those that showed alteration in at least one of the MSI-target genes that make up each gene function. MSI+ patients with

altered MSI-target genes related to gene functions: DNA repair, cell signaling, and adhesion/cytoskeleton were more likely to survive for five years when compared with MSI- patients (Supplementary Table IV).

Finally, based on the reported role of MSI in 5-fluorouracil (5-FU) and the high use of 5-FU and oxaliplatin in CRC treatment, we estimate the CSS stratified by MSI status for neoadjuvant and adjuvant protocols in cases with a clinical-stage II and III. We observed no impact in patient survival for any of the treatments used when comparing the MSI status (Supplementary Table V).

DISCUSSION

The current study performed the most extensive characterization of clinicopathological aspects from Brazilian CRC patients. Regarding MSI status, it analyzed MSI mutator phenotype in alterations of different MSI-target genes, and identified whether MSI status could influence patients' clinicopathological features and disease outcomes.

With the advent of personalized medicine, the need for and importance of prognostic or diagnostic biomarkers has increased. MSI status has been used as a biomarker for several purposes in CRC, such as (I) screening hereditary cases; (II) prognostic marker, where cases with MSI have a better prognosis than those who do not; (III) resistance to 5-FU therapy and sensitivity to irinotecan and recently, (IV) immunotherapy response.

The frequency of MSI+ among CRC varies in the literature, ranging from 6% to 20% [11, 19, 28–30]. According to TCGA (The Cancer Genome Atlas Network) data for CRC, 16% of analyzed samples showed MSI+ [31]. We observed a frequency of MSI+ in approximately 10% of sporadic CRC patients, like those reported in other studies. This discrepancy of MSI+ proportion among our results and other works can be caused by several factors, such as methodological differences used in MSI determination; inclusion of hereditary and sporadic CRC cases; distinct AJCC staging inclusion of patients; different patients' ethnicities, and environmental criteria that may affect the presence of MSI in CRC. The present work conclusive elucidates the MSI frequency by analyzing a large number of Brazilian CRC, originated from distinct Brazilian regions and representing all clinical stages.

The MSI status association with clinicopathological features showed significant associations in multivariate analyses, such as tumors in the right colon, histological grade III/undifferentiated, and clinical stage IV. These results are in accordance with international literature [11, 14, 29, 30, 32, 33], highlighting the

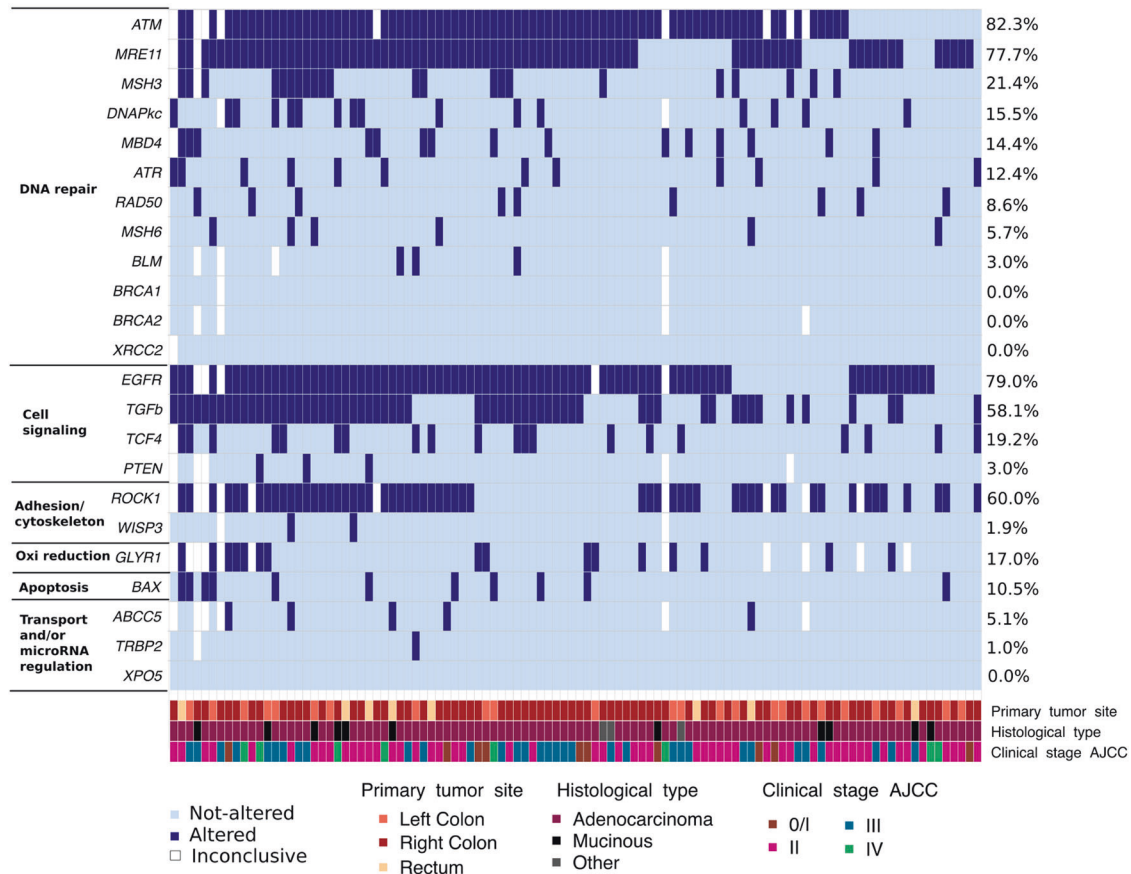


Fig. 2 MSI-target gene analyses. Upper panel: the columns represent the MSI+ patients, and the lines represent each analyzed microsatellite region of the MSI-target genes. At right, the percentage of altered cases are shown for contraction or expansion (dark blue) in the analyzed microsatellite region of each MSI-target gene stratified by gene function (at left). Light blue and white rectangles represent not-altered and inconclusive cases, respectively. The most altered genes are at the top of each gene function section. Lower panel: the columns represent the MSI+ patients analyzed for MSI-target gene, and the lines represent the distribution of clinicopathological features (primary tumor site, histological type, and clinical stage AJCC) for each sample.

distinct genetic relation about tumor location at colon and reinforcing the role of MSI as a prognostic biomarker in CRC.

The survival analysis of the present study showed a 5-years probability of 62.3% for CRC patients. When we stratified the cases regarding the MSI status (MSI- and MSI+), we observed a statistically significant difference between the two curves, with a 5-years probability of 60.5% for the MSI- and 77.7% for MSI+ patients. There are reports of significant differences between MSI statuses in the international literature, as Yoon et al. demonstrated that MSI+ patients had higher disease-free survival rates and overall survival than those MSI- [29]. Differently, there was no statistical difference between the survival curves when Nam et al. analyzed cases of advanced CRC stratified according to MSI status [33]. MSI status is often associated with survival in CRC. The meta-analysis by Guastadisegni et al. showed that MSI+ CRC was associated with a 40% higher survival rate than MSI- CRC [34]. Another study with >7000 cases reported that MSI+ patients have a significantly better prognosis than those with MSI- tumors [35].

Significant associations were observed between survival and MSI+ CRC patients' features, including better survival probability for patients with a clinical-stage II and III, absence of angiolymphatic and perineural invasions, and no recurrence of the disease. The clinical-stage variable showed a significant difference between the survival curves in both MSI statuses. Given the importance of staging as a prognostic factor, estimates of survival stratified by MSI status for stages 0/I, II, III, and IV were performed. In general, there was a statistically significant difference for clinical

stage III, where MSI+ patients showed a better 5-years probability than MSI- patients. These results are consistent with studies that state that MSI+ patients have a better prognosis when paired by stage than MSI- [36, 37]. However, this data is not consensual, and other studies had found no differences when patients in stages II and III were considered separate according to the location of the tumor [29, 38].

MSI status association with chemotherapy response is unclear. In the present study, we estimated survival stratified by MSI status and treatment based on 5-FU and oxaliplatin. There was no statistically significant difference for any of the analyzes. In a clinical trial with stage II-III colon cancer patients, those who were MSI+ had a better prognosis, but there was no association between MSI status and the benefits of chemotherapy [39]. The 5-FU treatment is used to treat CRC, being recommended in high-risk stage II cases and as a first-line for stages III and IV [40]. Several studies analyze the influence that MSI has on the response to 5-FU [37] and have described that MSI status is not a predictive biomarker of response to 5-FU [41, 42]. Webber et al. carried out a meta-analysis involving 9212 patients treated or not with 5-FU and concluded that the therapeutic regimen improved disease-free survival and overall survival, but the status of MSI did not influence the response to 5-FU-treatment [40]. Moreover, Alex et al. suggest that MSI+ phenotype is predictive of resistance to oxaliplatin-based chemotherapy, suggesting biological heterogeneity within the MSI+ CRC metastatic patients [43]. Currently, at Barretos Cancer Hospital, stage II MSI positive cases are not

Table 2. Kaplan–Meier estimates of cancer-specific survival for colorectal cancer patients and for MSI status ($n = 1002$).

Variable	Category	MSI status							
		MSI–			MSI+				
		Cases	Deaths	5 years survival (%)	p value*	Cases	Deaths	5 years survival (%)	p value*
Gender	Male	471	223	57.7	0.02	44	13	69.0	0.24
	Female	426	177	63.7		51	9	81.2	
Age at diagnoses (in years)	<50	261	122	62.4	0.39	29	11	62.1	0.09
	≥50 to <75	548	248	59.1		53	9	81.5	
	≥75	88	30	64.0		13	2	82.5	
Primary tumor site	Right colon	177	77	59.5	0.64	65	13	78.7	0.45
	Left colon	443	189	61.8		22	7	66.7	
	Rectum	277	134	59.2		8	2	71.4	
Clinical stage (AJCC)	0/I	115	15	93.4	<0.001	–	–	–	<0.001
	II	323	90	77.0		53	7	85.6	
	III	297	146	56.1		35	10	71.4	
	IV	160	149	13.7		7	5	16.7	
Histological type	Adenocarcinoma	846	378	60.8	0.003	82	17	78.1	0.29
	Mucinous	47	18	61.8		10	4	60.0	
	Others	4	4	zero		3	1	50.0	
Histological grade	I/II	845	366	62.1	<0.001	70	14	78.8	0.52
	III/Undifferentiated	40	29	24.1		23	6	71.8	
Angiolymphatic invasion	No	553	172	73.4	<0.001	58	8	85.2	0.03
	Yes	266	173	40.0		34	11	65.6	
Perineural invasion	No	647	237	67.8	<0.001	74	11	84.1	<0.001
	Yes	111	78	35.3		12	7	41.7	
Presence synchronous Tumors	No	848	383	60.3	0.44	87	20	75.7	0.73
	Yes	49	17	64.4		8	2	71.4	
Presence of recurrence	No	615	175	72.0	<0.001	74	7	89.8	<0.001
	Yes	282	225	37.4		21	15	28.6	
Neoadjuvant chemotherapy	No	16	10	42.9	0.02	2	zero	100	0.54
	Yes	153	69	62.5		6	2	66.7	
Adjuvant chemotherapy	No	483	248	51.9	<0.001	50	12	73.0	0.72
	Yes	401	147	70.5		41	10	75.6	
Radiotherapy	No	697	302	60.7	<0.001	86	20	75.4	0.72
	Yes, neoadjuvant	158	71	62.9		7	2	66.7	
	Yes, adjuvant	19	7	78.9		2	zero	100	
	Yes, palliative	21	20	19.0		–	–		

MSI–: cases without microsatellite instability. MSI+: cases with microsatellite instability. Significant associations are indicated in bold.

*Log Rank test.

Table 3. Multivariate analysis of cancer-specific survival associated with different clinicopathological characteristics and treatment of patients with colorectal cancer by MSI status.

MSI status	Variable	Category	HR	CI 95%	p value
MSI–	Gender	Male	1	–	–
		Female	0.79	0.63–0.98	0.036
	Clinical stage (AJCC)	0/I	1	–	–
		II	2.31	1.29–4.15	0.005
		III	4.01	2.20–7.29	<0.001
		IV	10.1	5.68–18.02	<0.001
	Angiolymphatic invasion	No	1	–	–
		Yes	1.67	1.32–2.12	<0.001
	Presence of recurrence	No	1	–	–
		Yes	2.99	2.39–3.75	<0.001
Adjuvant chemotherapy	No	1	–	–	
	Yes	0.59	0.44–0.78	<0.001	
MSI+	Age at diagnoses (in years)	<50	1	–	–
		≥50 to <75	0.30	0.10–0.87	0.028
		≥75	0.32	0.05–2.12	0.238
	Clinical stage (AJCC)	II	1	–	–
		III	1.26	0.41–3.91	0.687
		IV	5.15	1.22–21.71	0.026
	Perineural invasion	No	1	–	–
		Yes	2.96	1.01–8.69	0.047
	Presence of recurrence	No	1	–	–
Yes		8.89	3.06–25.85	<0.001	

MSI–: cases without microsatellite instability. MSI+: cases with microsatellite instability. Significant associations are indicated in bold. HR hazard ratio, CI confidential interval.

Table 4. Kaplan–Meier estimates of cancer-specific survival for Clinical stage (AJCC) of colorectal cancer patients by MSI status.

Clinical stage (AJCC)	MSI status	Cases	Deaths	5 years survival (%)	p value*
0/I	MSI–	115	15	93.4	0.29
	MSI+	10	zero	100.0	
II	MSI–	323	90	77.0	0.09
	MSI+	53	7	85.6	
III	MSI–	297	146	56.1	0.04
	MSI+	35	10	71.4	
IV	MSI–	160	149	13.7	0.53
	MSI+	7	5	16.7	

MSI–: cases without microsatellite instability. MSI+: cases with microsatellite instability. Significant associations are indicated in bold.

*Log Rank test.

treated with 5-FU-base chemotherapy, following NCCN (National Comprehensive Cancer Network) guidelines and Ribic et al. [44].

MSI phenotype may affect any microsatellite region across the genome leading to a characteristic mutation burden, and each altered microsatellite region contributes to MSI+ CRC heterogeneity. The present work was the first to analyze the frequency of MSI-target genes in a representative number of Brazilian CRC patients. We observed that the five most altered genes were: *ATM*, *EGFR*, *MRE11*, *ROCK1*, and *TGFBR11*, and the absence of alterations in *BRCA1*, *BRCA2*, *XPO5*, and *XRCC2*. Mutation in *MRE11*, namely at the microsatellite tract of 11(T) located at intron 4, is observed in approximately 80% of MSI tumors and leads to aberrant splicing and a truncated protein [45]. Interestingly, Vilar et al. showed that

MRE11 deficiency could increase sensitivity to PARP inhibitors [46]. Another reported MSI-target gene is the *TGFBR11*, reported mutated in 69–90% of cases [47, 48]. Inactivation of *TGFBR11* signaling is one of the steps in CRC progression. A possible predictive role of prognosis for *TGFBR11* has also been reported when observing that CRC patients with stage III, MSI+, and mutation in *TGFBR11* treated with chemotherapy based on 5-FU had a better prognosis [49, 50]. We also estimated the cancer-specific survival for MSI-target gene function by MSI status. We observed a statistically significant difference between the survival curves for MSI-target genes involved with gene functions: DNA repair, cellular signaling, and adhesion/cytoskeleton. MSI+ cases with altered MSI-target genes linked to these gene functions were

more likely to survive in 5 years when compared to MSI– cases. In general, these gene functions are crucial for tumors development/progression, genes present in them have been used for the development of new drugs that contribute to a better response and patient survival, and data about changes in these pathways are essential for guidance in clinical conduct, which demonstrates the presence and importance of genetic heterogeneity among MSI cases.

The genetic ancestry of the present cases was previously analyzed and great admixture in composition was observed: African 12.7% (SD = 15.7%), European 74.2% (SD = 20.6%), Eastern Asian 6.5% (SD = 11.3%), and Amerindian 6.6% (SD = 7.1%) [18]. In the present study, there was no correlation between the different ancestry proportions and MSI status and neither association of MSI status frequencies among Brazilian regions of origin from patients despite the divergence of ancestral components present in different country regions.

The present pioneering study determined the association of MSI status on 5-year survival and association of clinicopathological and molecular features in >1000 Brazilian CRC patients. We observed 10% of MSI+ frequency, tumors preferentially localized in the right colon, clinicopathological characteristics associated with less aggressiveness, and we observed a significant difference in the survival of these patients. MSI+ cases showed changes in several MSI-target genes, being the most altered related to functions like DNA repair, DNA damage sensor, and cellular signaling. The present study demonstrated the genetic heterogeneity present in MSI+ CRC patients and may contribute to the clinical management strategies of these patients.

DATA AVAILABILITY

Data supporting the results reported in this manuscript can be found within the article and its Supplementary Files.

REFERENCES

- Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA: Cancer J Clin.* 2021;71:7–33.
- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin.* 2021;71:209–49.
- Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA). Ministério da Saúde. Estimate/2020 – Cancer Incidence in Brazil. 2019.
- Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. *Nat Rev Dis Primers.* 2015;15065:1–25.
- Safiri S, Sepanlou SG, Ikuta KS, Bisignano C, Salimzadeh H, Delavari A, et al. The global, regional, and national burden of colorectal cancer and its attributable risk factors in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol.* 2019;4:913–33.
- de Carvalho T, Borges AK, da Silva I. Incidence of colorectal cancer in selected countries of Latin America: age-period-cohort effect. *Asian Pacific J Cancer Prev.* 2020;21:3421–8.
- Markowitz SD, Bertagnoli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. *N. Engl J Med* [Internet]. 2009;361:2449–60. <http://www.ncbi.nlm.nih.gov/pubmed/20018966>.
- Boland CR, Goel A. Microsatellite instability in colorectal cancer. *Gastroenterol* [Internet]. 2010;138:2073–e3. <http://www.ncbi.nlm.nih.gov/pubmed/20420947>.
- Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Song S, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* [Internet]. 2015;21:1350–6. <http://www.ncbi.nlm.nih.gov/pubmed/26457759>.
- Valle L, Vilar E, Tavtigian SV, Stoffel EM. Genetic predisposition to colorectal cancer: syndromes, genes, classification of genetic variants and implications for precision medicine. *J Pathol John Wiley Sons Ltd.* 2019;247:574–88.
- Vilar E, Gruber SB. Microsatellite instability in colorectal cancer: the stable evidence. *Nat Rev Clin Oncol.* 2010;7:153–62.
- Kim TM, Laird PW, Park PJ. The landscape of microsatellite instability in colorectal and endometrial cancer genomes. *Cell* [Internet]. 2013;155:858–68. <http://www.ncbi.nlm.nih.gov/pubmed/24209623>.
- Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med* [Internet]. 2016;22:1342–50. <http://www.ncbi.nlm.nih.gov/pubmed/27694933>.
- Li K, Luo H, Huang L, Luo H, Zhu X. Microsatellite instability: a review of what the oncologist should know. Vol. 20, *Cancer Cell International*. BioMed Central Ltd.; 2020.
- Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science.* 2017;357:409–13.
- Xicola RM, Gagnon M, Clark JR, Carroll T, Gao W, Fernandez C, et al. Excess of proximal microsatellite-stable colorectal cancer in African Americans from a multiethnic study. *Clin Cancer Res.* 2014;20:4962–70.
- Carethers JM, Murali B, Yang B, Doctolero RT, Tajima A, Basa R, et al. Influence of race on microsatellite instability and CD8+ T cell infiltration in colon cancer. *PLoS ONE.* 2014;9:e100461.
- Durães RO, Berardinelli GN, da Costa AM, Scapulatempo-Neto C, Pereira R, Oliveira MA, et al. Role of Genetic Ancestry in 1,002 Brazilian Colorectal Cancer Patients From Barretos Cancer Hospital. *Front Oncol.* 2020;10:1–12.
- Ashktorab H, Ahuja S, Kannan L, Llor X, Ellis NA, Xicola RM, et al. A meta-analysis of MSI frequency and race in colorectal cancer. *Oncotarget.* 2016;7:34546–57.
- dos Santos LV, Faria TMV, Lima ABC, Abdalla KC, de Moraes ED, Cruz MR, et al. Timing of adjuvant chemotherapy in colorectal cancer. *Colorectal Dis.* 2016;18:871–6.
- de Paula AE, Galvão H de CR, Bonatelli M, Sabato C, Fernandes GC, et al. Clinicopathological and molecular characterization of Brazilian families at risk for Lynch syndrome. *Cancer Genetics.* 2021;254–255:82–91.
- Berardinelli GN, Scapulatempo-Neto C, Durães R, Antônio de Oliveira M, Guimarães D, Reis RM. Advantage of HSP110 (T17) marker inclusion for microsatellite instability (MSI) detection in colorectal cancer patients. *Oncotarget* [Internet]. 2018;9:28691–701. <https://www.oncotarget.com/lookup/doi/10.18632/oncotarget.25611>.
- Campanella NC, Berardinelli GN, Scapulatempo-Neto C, Viana D, Palmero EI, Pereira R, et al. Optimization of a pentaplex panel for MSI analysis without control DNA in a Brazilian population: correlation with ancestry markers. *Eur J Hum Genet* [Internet]. 2014;22:875–80. <http://www.ncbi.nlm.nih.gov/pubmed/24193342>.
- Viana-Pereira M, Almeida I, Sousa S, Mahler-Araújo B, Seruca R, Pimentel J, et al. Analysis of microsatellite instability in medulloblastoma. *Neuro Oncol* [Internet]. 2009;11:458–67. <http://www.ncbi.nlm.nih.gov/pubmed/19179424>.
- Pereira R, Phillips C, Pinto N, Santos C, dos Santos SE, Amorim A, et al. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. *PLoS ONE* [Internet]. 2012;7:e29684. <http://www.ncbi.nlm.nih.gov/pubmed/22272242>.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genet* [Internet]. 2000;155:945–59. <http://www.stats.ox.ac.uk/pritch/home.html>.
- Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genet* [Internet]. 2003;164:1567–87.
- Advani SM, Shi Q, Overman MJ, Loree JM, Lam M, Morris V, et al. Patient-reported symptom outcomes and microsatellite instability in patients with metastatic colorectal cancer. *Clin Colorectal Cancer.* 2020;19:48–56.e2.
- Yoon YS, Yu CS, Kim TW, Kim JH, Jang SJ, Cho DH, et al. Mismatch repair status in sporadic colorectal cancer: immunohistochemistry and microsatellite instability analyses. *J Gastroenterol Hepatol* [Internet]. 2011;26:1733–9. <http://www.ncbi.nlm.nih.gov/pubmed/21615788>.
- Corso G, Pascale V, Flauti G, Ferrara F, Marrelli D, Roviello F. Oncogenic mutations and microsatellite instability phenotype predict specific anatomical subsite in colorectal cancer patients. *Eur J Hum Genet.* 2013;21:1383–8.
- Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. *Nat* [Internet]. 2012;487:330–7. <http://www.ncbi.nlm.nih.gov/pubmed/22810696>.
- Leite SM, Gomes KB, Pardini VC, Ferreira AC, Oliveira VC, Cruz GM. Assessment of microsatellite instability in colorectal cancer patients from Brazil. *Mol Biol Rep.* [Internet]. 2010;37:375–80. <http://www.ncbi.nlm.nih.gov/pubmed/19784864>.
- Nam SK, Yun S, Koh J, Kwak Y, Seo AN, Park KU, et al. BRAF, PIK3CA, and HER2 oncogenic alterations according to KRAS mutation status in advanced colorectal cancers with distant metastasis. *PLoS ONE* [Internet]. 2016;11:e0151865. <http://www.ncbi.nlm.nih.gov/pubmed/26991109>.
- Guastadisegni C, Colafranceschi M, Ottini L, Dogliotti E. Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data. *Eur J Cancer* [Internet]. 2010;46:2788–98. <http://www.ncbi.nlm.nih.gov/pubmed/20627535>.
- Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. *J Clin Oncol* [Internet]. 2005;23:609–18. <http://www.ncbi.nlm.nih.gov/pubmed/15659508>.

36. Laghi L, Malesci A. Microsatellite instability and therapeutic consequences in colorectal cancer. *Dig Dis* [Internet]. 2012;30:304–9. <http://www.ncbi.nlm.nih.gov/pubmed/22722556>.
37. Kocarnik JM, Shiovitz S, Phipps AI. Molecular phenotypes of colorectal cancer and potential clinical applications. *Gastroenterol Rep. (Oxf)* [Internet]. 2015;3:269–76. <http://www.ncbi.nlm.nih.gov/pubmed/26337942>.
38. Ronnekleiv-Kelly SM, Burkhart RA, Pawlik TM. Molecular markers of prognosis and therapeutic targets in metastatic colorectal cancer. *Surg Oncol* [Internet]. 2016;25:190–9. <http://www.ncbi.nlm.nih.gov/pubmed/27566022>.
39. Gavin PG, Colangelo LH, Fumagalli D, Tanaka N, Remillard MY, Yothers G, et al. Mutation profiling and microsatellite instability in stage II and III colon cancer: an assessment of their prognostic and oxaliplatin predictive value. *Clin Cancer Res* [Internet]. 2012;18:6531–41. <http://www.ncbi.nlm.nih.gov/pubmed/23045248>.
40. Webber EM, Kauffman TL, O'Connor E, Goddard KA. Systematic review of the predictive effect of MSI status in colorectal cancer patients undergoing 5FU-based chemotherapy. *BMC Cancer* [Internet]. 2015;15:156. <http://www.ncbi.nlm.nih.gov/pubmed/25884995>.
41. Sturgeon CM, Duffy MJ, Stenman UH, Lilja H, Brunner N, Chan DW, et al. National Academy of Clinical Biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers. *Clin Chem* [Internet]. 2008;54:e11–79. <http://www.ncbi.nlm.nih.gov/pubmed/19042984>.
42. Duffy MJ, van Dalen A, Haglund C, Hansson L, Holinski-Feder E, Klapdor R, et al. Tumour markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines for clinical use. *Eur J Cancer* [Internet]. 2007;43:1348–60. <http://www.ncbi.nlm.nih.gov/pubmed/17512720>.
43. Alex AK, Siqueira S, Coudry R, Santos J, Alves M, Hoff PM, et al. Response to chemotherapy and prognosis in metastatic colorectal cancer with DNA deficient mismatch repair. *Clin Colorectal Cancer*. 2017;16:228–39.
44. Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM, et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N. Engl J Med*. 2003;349:247–57.
45. Giannini G, Ristori E, Cerignoli F, Rinaldi C, Zani M, Viel A, et al. Human MRE11 is inactivated in mismatch repair-deficient cancers. *EMBO Rep*. 2002;3:248–54.
46. Vilar E, Bartnik CM, Stenzel SL, Raskin L, Ahn J, Moreno V, et al. MRE11 deficiency increases sensitivity to Poly(ADP-ribose) polymerase inhibition in microsatellite unstable colorectal cancers. *Cancer Res*. 2011;71:2632–42.
47. Grady WM, Markowitz SD. TGF- β signaling pathway and tumor suppression. In: Derynck KR, editor. *The TGF- β family*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2008. p. 889–938.
48. Tsang AH, Cheng KH, Wong AS, Ng SS, Ma BB, Chan CM, et al. Current and future molecular diagnostics in colorectal cancer and colorectal adenoma. *World J Gastroenterol* [Internet]. 2014;20:3847–57. <http://www.ncbi.nlm.nih.gov/pubmed/24744577>.
49. Pawlik TM, Raut CP, Rodriguez-Bigas MA. Colorectal carcinogenesis: MSI-H versus MSI-L. *Dis Markers* [Internet]. 2004;20:199–206. <http://www.ncbi.nlm.nih.gov/pubmed/15528785>.
50. Watanabe T, Wu TT, Catalano PJ, Ueki T, Satriano R, Haller DG, et al. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. *N. Engl J Med* [Internet]. 2001;344:1196–206. <http://www.ncbi.nlm.nih.gov/pubmed/11309634>.

ACKNOWLEDGEMENTS

The authors acknowledge Dr. Adriane Feijó and Dr. Luciane Sussuchi for their assistance in preparing the figures.

AUTHOR CONTRIBUTIONS

GNB participated in study design, collection of the data, analyzed the data, and prepared the manuscript. RD participated in the collection of the clinical data and analyzed the data. AMC participated in the collection of the data and analyzed the data. AB participated in re-evaluation of the clinical data and analyzed the data. MO participated in the statistical analyses and analyzed the data. RP participated in the ancestry analyses and analyzed the data. CS-N participated in the collection of the data and re-evaluation of the diagnoses. DPG participated in re-evaluation of the diagnoses and analyzed the data. RMR participated in study design, coordination, analyses of the data, and prepared the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was partially supported by Barretos Cancer Hospital Internal Research Funds (PAIP) of participant authors and the Public Ministry of Labor – Campinas, São Paulo (Research, Prevention, and Education of Occupational Cancer). RMR was the recipient of a National Council of Technological and Scientific Development (CNPq) scholarship. AMC had financial support from the São Paulo Government (São Paulo Research Foundation – FAPESP) through the grants numbers 2019/21722–0 and 2018/22097–0.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study was evaluated and approved by Institutional Ethics Committee (protocol number: 600/2012 - CAAE: 02468812.30000.5437).

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41431-022-01104-y>.

Correspondence and requests for materials should be addressed to Rui Manuel Reis.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.