

Epidermal growth factor receptor intron 1 polymorphism and microsatellite instability in sporadic colorectal cancer

SONJA MARINOVIĆ^{1*}, KRISTINA VUKOVIĆ^{1*}, ANITA ŠKRTIĆ², MIRKO POLJAK³,
SARA PETEK¹, LARA PETEK¹ and SANJA KAPITANOVIĆ¹

¹Laboratory for Personalized Medicine, Division of Molecular Medicine, Ruder Boskovic Institute;
Departments of ²Pathology and ³Surgery, Clinical Hospital Merkur, 10000 Zagreb, Croatia

Received June 18, 2020; Accepted November 9, 2020

DOI: 10.3892/ol.2020.12392

Abstract. Epidermal growth factor receptor (EGFR) expression is commonly upregulated in sporadic colorectal cancer (CRC) and its high expression is associated with poor prognosis in patients with CRC. CA-SSR1 is a dinucleotide CA repeat of the *EGFR* gene that can modulate *EGFR* transcription and is a potential target of the mismatch repair machinery in tumours with microsatellite instability (MSI). In the present study, 160 sporadic colon cancer samples were analysed for *EGFR* CA-SSR1 polymorphism and MSI status. Additionally, *EGFR* mRNA and protein expression levels in the tumour centre and in the invasive tumour front, compared with those in adjacent normal tissue samples, were evaluated in 80 tumour samples. An inverse association was identified between *EGFR* mRNA levels and the sum of repeats in both alleles of the CA-SSR1 polymorphism in normal tissues. Changes in CA-SSR1 were detected in the tumour centre as well as in the invasive tumour front and metastases in all MSI high (MSI-H) tumours. Analysis of EGFR expression at the mRNA and protein levels according to MSI status revealed lower *EGFR* mRNA and protein expression in MSI-H tumours than microsatellite-stable (MSS) tumours. Furthermore, higher EGFR levels in the invasive tumour front compared with in the tumour centre in MSS tumours were identified, suggesting a role of EGFR in tumour progression and higher invasive potential of MSS than MSI-H tumours.

Introduction

Sporadic colorectal cancer (CRC) is the second most common cause of cancer death in developed countries with an estimated

1.8 million new cases and 862,000 deaths in 2018 (1,2). Despite the overall advances in diagnosis and therapy, survival rates for colorectal cancer remain disappointing at approximately 65% depending on the stage. Hence, finding molecular markers that can improve patient diagnosis and treatment selection is necessary.

Epidermal growth factor receptor (EGFR) is a membrane receptor of the receptor tyrosine kinase (ErbBs) family that plays an important role in cell proliferation, survival, differentiation and invasion in sporadic colon cancer (3,4). Several studies have shown that *EGFR* is overexpressed in approximately 50% of colon tumours (5-7) and there is an increased level of EGFR protein at the invasive tumour front (ITF) in comparison to the tumour centre (8,9). This increased EGFR expression has been related to the presence of tumour budding leading to tumour invasion and other aggressive pathohistological features (10-13). Overexpression of EGFR protein in CRC can be rarely attributed to gene amplification and is more often attributed to polymorphisms in the *EGFR* gene (14).

Polymorphic regions are excellent candidates as possible prognostic biomarkers for cancer patients. Their main advantage is that they can be easily assessed from blood and normal and tumour tissue and determined by straightforward and well-established methods. One of these regions is the CA dinucleotide repeat polymorphism in intron 1 of the *EGFR* gene. CA-SSR1 is important due to its close location to the second enhancer (15) which endows it with the ability to influence the expression of the *EGFR* gene (16-21). Given that CA-SSR1 could modulate *EGFR* transcription, changes in its sequence could also alter the levels of EGFR protein. Different studies investigated the role of CA repeats and its prognostic implication in various types of cancer and some of them have brought into connection an increasing number of repeats and decreasing levels of both EGFR mRNA and protein expression (22,23). However, the results and its potential predictive impact in CRC remain contrasting and inconclusive (24,25).

Sporadic CRCs can be divided into microsatellite-stable (MSS) tumours and tumours having microsatellite instability (MSI) as a result of failure in the DNA mismatch repair (MMR) system (26). This failure results in changes in the length of microsatellite sequences, potentially also affecting the *EGFR* CA-SSR1 polymorphism. Even though it is recognized that MSI can affect repeat elements of the *EGFR* gene,

Correspondence to: Professor Sanja Kapitanović, Laboratory for Personalized Medicine, Division of Molecular Medicine, Ruder Boskovic Institute, Bijenicka cesta 54, 10000 Zagreb, Croatia
E-mail: kapitan@irb.hr

*Contributed equally

Key words: sporadic colorectal cancer, epidermal growth factor receptor, CA-SSR1 polymorphism, microsatellite instability

and subsequent *EGFR* expression (27,28), clinical and pathological significance is still not extensively studied.

Therefore, we decided to investigate the effect of the *EGFR* CA-SSR1 polymorphism on mRNA and protein expression by considering microsatellite status in CRC tumours. Furthermore, we wanted to examine the difference in *EGFR* mRNA and protein expression between the tumour centre and invasive tumour front in accordance to MSI status to clarify the role of EGFR in CRC tumour progression. Additionally, we performed a correlation analysis between the EGFR expression and CRC clinicopathological characteristics in MSS and MSI-H tumours.

Materials and methods

Study subjects and DNA isolation. Our study included tumour samples from 160 patients diagnosed with sporadic colon adenocarcinoma obtained from the Croatian Tumour Bank (29). Tissue samples were collected from 2013 to 2019 during routine surgery performed in Merkur Clinical Hospital. Fresh tumour samples were stored at -80°C until DNA and RNA extraction. From each patient two samples of tumour were obtained: One corresponding to the tumour centre (T1) and the other corresponding to the invasive tumour front (T2) as well as adjacent normal colon tissue. If metastasis (M) was present, samples were retrieved for the analysis. In MSS tumours 15 and in MSI tumours 2 metastasis samples were obtained. Before use in the study, each specimen was verified by a pathologist (A.Š.).

DNA was extracted from the blood and tumour tissues as well as corresponding normal tissue samples located 15 cm from the tumour edge. DNA extraction was performed using proteinase K digestion and phenol-chloroform extraction (30).

PCR. For MSI analysis paired normal and tumour DNA was analysed for changes in five loci each, using previously published Bethesda panel (26).

The primer sequences used were as follows: D2S123: forward, 5'-AAACAGGATGCCTGCCTTTA-3' and reverse, 5'-GGACTTTCCACCTATGGGAC-3'; D5S346: forward, 5'-ACTCACTCTAGTGATAAATCGGG-3' and reverse, 5'-AGCAGATAAGACAGTATTACTAGTT-3'; D17S250: forward, 5'-GGAAGAATCAAATAGACAAT-3' and reverse, 5'-GCTGGCCATATATATATTTAAACC-3'; BAT-25a: forward, 5'-TCGCCTCCAAGAATGTAAGT-3' and reverse, 5'-TCTGCATTTTAACTATGGCTC-3'; BAT-26: forward, 5'-CTGCGTAATCAAGTTTTTAG-3' and reverse, 5'-AACCATCAACATTTTAAACCC-3'. Samples were considered MSI low (MSI-L) samples if only one marker was changed and MSI high (MSI-H) if at least two out of five markers showed instability.

For *EGFR* intron 1 polymorphism genotyping, forward, 5'-GGGCTCACAGCAAACCTTCTC-3' and reverse, 5'-AAGCCAGACTCGCTCATGTT-3' *EGFR* primers were used.

Genomic DNA (100 ng) was used as a template in a reaction volume of 25 μl containing 5 pmol of each primer, 50 μM of each dNTP, and 1 U of Taq Gold DNA polymerase (Applied Biosystems; Thermo Fisher Scientific, Inc.). PCR tests were carried out in an Applied Biosystems GeneAmp PCR System 2400 for 30 cycles. Annealing temperatures for each primer

set were optimized in pilot studies before processing experimental samples.

Short tandem repeats analysis. Polymorphic marker analysis of D2S123, D5S346, D17S250 and *EGFR* intron 1 polymorphism was performed by non-denaturing polyacrylamide electrophoresis as previously described (31). In brief, the PCR product was mixed with loading buffer and loaded onto a 10% non-denaturing polyacrylamide gel. Electrophoresis was performed for 16 h at room temperature and gels were silver stained. MSI was defined as a visible change in the allele:allele ratio in tumours compared with matching normal tissue.

Analysis of BAT-25 and BAT-26, as well as validation of fragment length for the *EGFR* intron 1 polymorphism, was carried out with an ABI Prism[®] 310 genetic analyser with a GeneScan Analyzer (Life Technologies; Thermo Fisher Scientific, Inc.). The primers used for the genetic analyses were labelled with fluorescent dye. The GS500 ROX (-250 LIZ; Thermo Fisher Scientific, Inc.) size marker was added to each sample.

RNA extraction, cDNA synthesis and quantitative PCR (qPCR). Total RNA was extracted from snap frozen samples of resected colon carcinoma and corresponding normal tissues using TRIzol reagent (Invitrogen; Thermo Fisher Scientific, Inc.). To quantify *EGFR* expression levels, 1 mg of RNA was used to synthesize equal amounts of cDNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems; Thermo Fisher Scientific, Inc.). qPCR analysis was performed using an ABI PRISM 7300 sequence detection system (Applied Biosystems; Thermo Fisher Scientific, Inc.) and predeveloped TaqMan assay reagents: Hs01076090 for the *EGFR* gene and Hs01060665 for the beta-actin gene as an internal control. PCR was carried out according to the manufacturer's protocol. The method used for analysis was $2^{-\Delta\Delta C_q}$ (32).

Immunohistochemical (IHC) analysis of EGFR expression. Formalin-fixed and paraffin-embedded blocks were cut into 2 μm sections at least 24 h prior to immunostaining and mounted on microscope slides. Antigen retrieval was carried out in citrate buffer at 95°C for 5 min. Sections were incubated overnight with anti-EGFR monoclonal antibody (1:100, Santa Cruz Biotechnology). Secondary detection was performed using a goat anti-rabbit IgG antibody conjugated to horseradish peroxidase (HRP) (ready to use; Cell Signalling Technology). Antigen-antibody complexes were visualized by incubation in diaminobenzidine tetrahydrochloride (Dako) for 3 min and counterstained with haematoxylin. The EGFR stained slides were evaluated for staining intensity which was scored from 0 (no staining) to 3 (strongest staining) using light microscope at the magnification of $\times 100$. All control slides (without primary antibody staining) were negative for staining. For further analysis, EGFR expression was divided into a high and low category. EGFR expression was defined low when there was either no staining (Score 0) or there was a weak positive (light brown) staining (Score 1), and high EGFR expression was defined when staining intensity was either intermediate (Score 2) or strong (Score 3).

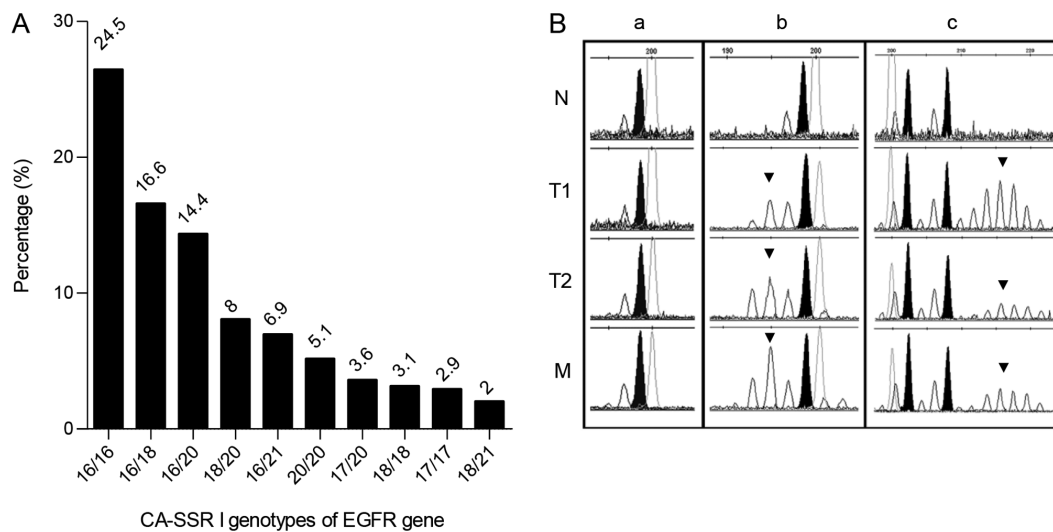


Figure 1. CA-SSR1 polymorphism of the *EGFR* gene and microsatellite instability. (A) Distribution of the 10 most common allele frequencies of *EGFR* intron1 polymorphism (CA-SSR1). (B) Electropherogram of the *EGFR* CA-SSR1 fragment length analysis carried out with a GeneScan Analyzer in N, T1, T2 and M samples from (a) MSS and (b and c) MSI high tumours. CA-SSR1 analysis revealed (a and b) homozygosity and (c) heterozygosity in the N. (a) MSS tumour exhibited no instability in all samples of the same patient. (b and c) MSI tumours exhibited instability with (b) additional shortened or (c) elongated alleles (indicated by arrows) in T1, T2 and M samples. Black peaks indicate germline alleles; the upper axis shows fragment sizes. N, T1, T2 and M samples were all collected from the same individual. EGFR, epidermal growth factor receptor; N, adjacent normal tissue; T1, tumour centre; T2, invasive tumour front; M, metastasis; MSS, microsatellite-stable; MSI, microsatellite instability.

Statistical analysis. Statistical analyses were performed using the GraphPad Prism statistical package (GraphPad Software, Inc.). Correlations between the *EGFR* CA-SSR1 genotype and *EGFR* mRNA expression were analysed using Spearman's correlation coefficient and linear regression analysis. The relationship between the sum of repeats in both alleles in the *EGFR* CA-SSR1 polymorphism and EGFR protein levels was examined by two-tailed unpaired Student's t-test and the Mann-Whitney test. Two-way analysis of variance (ANOVA) with Bonferroni correction was used to compare *EGFR* mRNA expression between the groups. For EGFR IHC analysis and further correlation with clinicopathological characteristics, contingency table with Fisher's exact test was used to calculate statistical significance. Overall survival rate was determined by Kaplan-Meier, and statistical differences between groups were calculated with the log-rank test. Data are presented as the mean \pm SEM. Values of * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$ were considered statistically significant.

Results

CA-SSR1 polymorphism of the *EGFR* gene and microsatellite instability. The *EGFR* intron 1 CA repeat polymorphism (CA-SSR1) was genotyped in 160 blood samples from patients with sporadic colon cancer. The number of CA repeats ranged in length from 15 to 21 and homozygous 16/16 CA repeats were the most frequent genotype (24.5%), followed by heterozygous 16/18 (16.6%) and 16/20 (14.4%) CA repeats. The frequencies of 10 most common genotypes (present in 140 patients; 87%) are shown in Fig. 1A.

Furthermore, tumour centre (T1), invasive tumour front (T2) and, if present, corresponding metastasis (M) samples in comparison to corresponding normal tissues (N) were analysed for MSI status and the *EGFR* CA-SSR1 polymorphism. Analysis showed that MSI was present in

Table I. MSI status of tumours.

Number of markers exhibiting instability	Number (%)	Interpretation
≥ 2	20 (12.5)	MSI-H
1	13 (8.1)	MSI-L
0	127 (79.4)	MSS

MSI, microsatellite instability; MSI-H, microsatellite instability high; MSI-L, microsatellite instability low; MSS, microsatellite stable.

33 (20.6%) tumour samples, 13 (8.1%) tumours were MSI-L and 20 of 160 analysed tumours (12.5%) were MSI-H. The remaining 127 (79.4%) tumours were classified as microsatellite stable (MSS) (Table I). CA-SSR1 polymorphism fragment length analysis in MSS colorectal tumours showed the same genotype in T1, T2, M and corresponding N tissue, in all samples (Fig. 1B-a). In contrast, in all MSI-H samples instability was detected in T1, T2 and M samples as additional shortening or elongation of alleles of the CA-SSR1 polymorphism (Fig. 1B-b and c).

Association of CA-SSR1 with *EGFR* mRNA and protein expression. To determine whether the number of CA-SSR1 CA repeats is associated with changes in *EGFR* transcription and protein expression levels, 80 specimens with the most frequent CA-SSR1 genotypes were analysed for *EGFR* mRNA and protein expression with regards to MSI status.

When we calculated the sum of CA repeats from both alleles per patient, the median sum of all samples present in study was 34 (range 31-42), with 52.2% patients having less than 34 repeats. Therefore, we classified patients as having

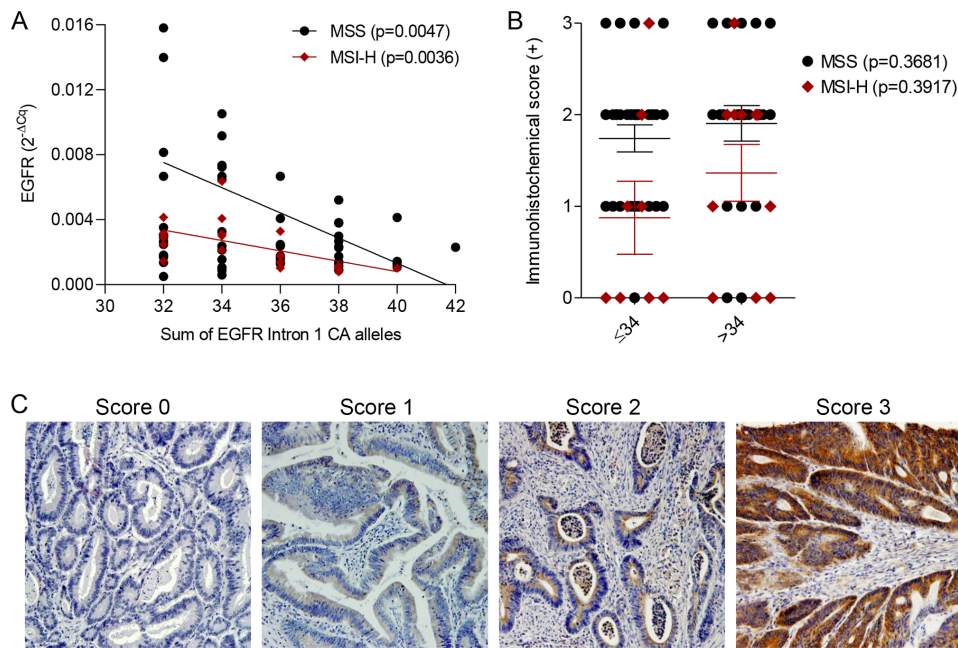


Figure 2. Association of CA-SSR1 with EGFR mRNA and protein expression in MSS and MSI-H sporadic colorectal tumours. (A) Spearman's correlation coefficient of the sum of CA repeats from both alleles per patient of the *EGFR* intron 1 polymorphism and *EGFR* mRNA expression levels in MSS ($r=-0.3576$; $P=0.0047$) and MSI-H ($r=-0.6335$; $P=0.0036$) tumours. (B) Correlation between the sum of CA repeats from both alleles per patient (≤ 34 or >34) of the CA-SSR1 polymorphism and EGFR immunohistochemical score. Statistical significance was analysed using (A) Spearman's correlation coefficient and linear regression analysis and (B) two-tailed unpaired Student's t-test. (C) Representative immunohistochemical images of epidermal growth factor receptor staining. Images represent the tumour staining scores between 0 and 3 (magnification, $\times 100$). EGFR, epidermal growth factor receptor; MSS, microsatellite-stable; MSI-H, microsatellite instability high.

either low (≤ 34) or high (>34) numbers of CA repeats in both alleles. An inverse correlation was found between a higher sum of CA repeats in the *EGFR* CA-SSR1 polymorphism and lower *EGFR* mRNA expression in both MSS ($P=0.0047$) and MSI-H sporadic colorectal tumours ($P=0.0036$) (Fig. 2A). However, there was no correlation between EGFR protein expression and CA-SSR1 genotype in MSS ($P=0.3681$) or MSI-H ($P=0.3917$) tumours (Fig. 2B). Fig. 2C shows representative immunohistochemical staining for EGFR in the tissues histologically classified as adenocarcinoma. Immunohistochemical staining was scored as negative (Score 0), weak (Score 1), intermediate (Score 2) or strong (Score 3) at a magnification of $\times 100$.

EGFR expression in the tumour centre and at the invasive margin. Since it is known that EGFR can promote cell migration, invasion, and metastatic dissemination we also evaluated *EGFR* mRNA and IHC protein levels in the tumour centre (T1) and invasive tumour front (T2) samples.

The analysis showed that *EGFR* mRNA levels were significantly higher in both T1 ($P=0.009$) and T2 ($P=0.024$) samples of MSS tumours than in normal tissues. However, in MSI-H tumours, *EGFR* mRNA levels were not significantly increased in either T1 or T2 samples in comparison to adjacent normal tissue samples ($P>0.999$). Both T1 and T2 of MSI-H tumours showed decreased *EGFR* mRNA levels in comparison to those in MSS tumour, nevertheless, this was not statistically significant. There was no difference in the expression of *EGFR* mRNA between normal tissues adjacent to MSS and normal tissues adjacent to MSI-H tumours (Fig. 3).

Immunohistochemical analysis further confirmed these results. EGFR protein expression was detected in 6 (35.5%)

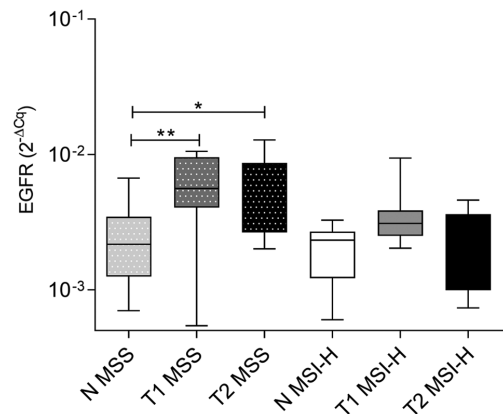


Figure 3. EGFR expression in the tumour centre and at the invasive tumour front. mRNA expression in N, T1 and T2 samples from MSS and MSI-H tumours. Statistical significance was analysed using two-way ANOVA. The data are presented as the mean \pm SEM. * $P<0.05$ and ** $P<0.01$. EGFR, epidermal growth factor receptor; N, adjacent normal tissue; T1, tumour centre; T2, invasive tumour front; MSS, microsatellite-stable; MSI-H, microsatellite instability high.

adjacent normal tissues, 50 (92.6%) T1 tissues and in 39 (97.5%) T2 tissues of MSS tumours as well as in 2 (28.5%) adjacent normal tissues, 12 (66.7%) T1 tissues and 6 (54.5%) T2 tissues of MSI-H tumours.

EGFR protein expression was significantly increased in tumour tissues in comparison to adjacent normal tissues in MSS tumours ($P<0.0001$). Moreover, the analysis showed a difference in EGFR protein expression between T1 and T2 samples of MSS tumours ($P=0.001$) (Table II and Fig. 4).

Table II. EGFR protein immunohistochemical expression in N, T1 and T2 samples from MSS and MSI-H tumours.

MSI status	EGFR immunohistochemical score (+)				P-values	
	0, n (%)	1, n (%)	2, n (%)	3, n (%)	N/T1 or T2	T1/T2
MSS						
N	11 (64.7)	5 (29.4)	1 (5.9)	0 (0)		
T1	4 (7.4)	12 (22.2)	24 (44.5)	14 (25.9)		0.001 ^e
T2	1 (2.5)	2 (5.0)	15 (37.5)	22 (55.0)	<0.0001 ^a <0.0001 ^b	0.799 ^g 0.040 ^h 0.001 ⁱ
MSI-H						
N	5 (71.4)	2 (28.6)	0 (0)	0 (0)		
T1	6 (33.3)	4 (22.3)	6 (33.3)	2 (11.1)	0.182 ^c	0.466 ^f
T2	5 (45.5)	1 (9.1)	2 (18.2)	3 (27.2)	0.197 ^d	

Letters indicate the difference between ^aMSS N and MSS T1, ^bMSS N and MSS T2, ^cMSI-H N and MSI-H T1, ^dMSI-H N and MSI-H T2, ^eMSS T1 and MSS T2, ^fMSI-H T1 and MSI-H T2, ^gMSS N and MSI-H N, ^hMSS T1 and MSI-H T1 and ⁱMSS T2 and MSI-H T2. P-values were obtained by Fisher's exact test. N, adjacent normal tissue; T1 tumour centre; T2 invasive tumour front; EGFR, epidermal growth factor receptor; MSI, microsatellite instability; MSS, microsatellite stable; MSI-H, microsatellite instability-high.

There was also significantly higher EGFR expression in MSS tumours than in MSI-H tumours in both T1 and T2 samples (P=0.040, P=0.001, respectively) (Table II and Fig. 4). However, there was no difference in the expression of EGFR protein in tumour tissues in comparison to adjacent normal tissues in MSI-H tumours or between normal tissues adjacent to MSS and normal tissues adjacent to MSI-H tumours (Table II and Figs. 4, S1 and S2). Immunohistochemical staining showed similar EGFR protein expression in T2 and liver or lymph node metastasis samples (Figs. 4, S1 and S2).

There was no correlation between EGFR protein expression and tumour size, histological grade, Dukes' stage of tumours (P>0.05) (Table III) in either MSS or MSI-H tumours. The survival of the two MSS tumour subgroups, based on EGFR immunohistochemical score, had no statistically significant difference (P=0.717) (Table IV).

Discussion

EGFR receptor overexpression is found in a wide range of cancers, including CRC, and it is associated with an aggressive tumour phenotype and poor prognosis (33). Nonetheless, the mechanisms regulating the levels of EGFR expression in cancer have not been fully characterized. The results of our study show that the CA-SSR1 of the *EGFR* gene is altered in MSI-H sporadic colorectal tumours and that it has an effect on EGFR expression at the mRNA level but not at the protein level in both MSS and MSI-H tumours. Moreover, we demonstrated that MSI-H tumours have lower *EGFR* mRNA and protein levels in the tumour centre and invasive tumour front than MSS tumours. Additionally, we confirmed that in MSS tumours EGFR expression is higher in the invasive tumour front than in the tumour centre.

In recent years, *EGFR* gene intron 1 length has been considered a factor affecting expression through modification of EGFR transcription. In this regard, it has been suggested that CA-SSR1 has an effect on gene transcription. This hypothesis was tested in several cell lines (20,22,23) as well as in head and neck, lung, pancreas, colon and mammary tumours (17-20,34,35), but the results were inconsistent and somewhat contradictory, most likely due to the limited number of analysed samples. We found the same distribution of CA-SSR1 alleles in normal tissues as previously reported (18,28,36), and the most common genotype was 16/16.

The distribution of CA-SSR1 alleles and MSI status, were analysed in normal and tumour tissues (tumour centre and invasive tumour front) as well as in corresponding metastasis samples when available. Our results showed for the first time that all MSI-H tumours showed instability in the CA-SSR1 polymorphism. However, the genotype varied from shortening to elongation, with presence of one additional allele to several of them, regardless of the length of the alleles present in normal tissues. CA-SSR1 instability was present in the tumour centre as well as in invasive tumour front and metastasis samples. To further characterize the possible effect of the CA-SSR1 polymorphism of the *EGFR* gene, and its changes in MSI-H tumours, we measured *EGFR* mRNA and protein levels. Since there is a lack of consensus regarding cut off values defining shorter versus longer CA repeats (37-40), we decided to take the median sum of the CA repeats from both alleles

Table III. Clinicopathological characteristics of 147 patients with sporadic adenocarcinoma stratified by MSI status.

Characteristics	MSI status						P-value
	MSS, n=127 (86.4%)			MSI-H, n=20 (13.6%)			
	Low EGFR expression, n (%)	High EGFR expression, n (%)		Low EGFR expression, n (%)	High EGFR expression, n (%)		
Tumor size, cm							0.3239
<5	11 (42.3)	15 (57.7)		3 (50.0)	3 (50.0)		
≥5	7 (26.9)	19 (73.1)		8 (61.5)	5 (38.5)		
Histological grade							0.4226
Well (1)	8 (40.0)	12 (60.0)		2 (50.0)	2 (50.0)		
Moderate and poor (2 and 3)	9 (74.3)	26 (25.7)		8 (53.3)	7 (46.7)		
Dukes' stage							0.5305
A+B	5 (21.7)	18 (78.3)		7 (70.0)	3 (30.0)		
C+D	9 (30.0)	21 (70.0)		4 (44.4)	5 (55.6)		

Low EGFR, immunohistochemical score 0 and 1; high EGFR, immunohistochemical score 2 and 3. MSI, microsatellite instability; MSS, microsatellite stable; MSI-H, microsatellite instability-high; EGFR, epidermal growth factor receptor.

Table IV. Survival of patients with microsatellite stable sporadic adenocarcinoma stratified by EGFR expression.

EGFR expression	Cases, n	Deaths, n	Mean survival, months	P-value
Low	5	3	21.6	0.7171
High	11	5	39.4	

Low EGFR, immunohistochemical score 0 and 1; high EGFR, immunohistochemical score 2 and 3. EGFR, epidermal growth factor receptor.

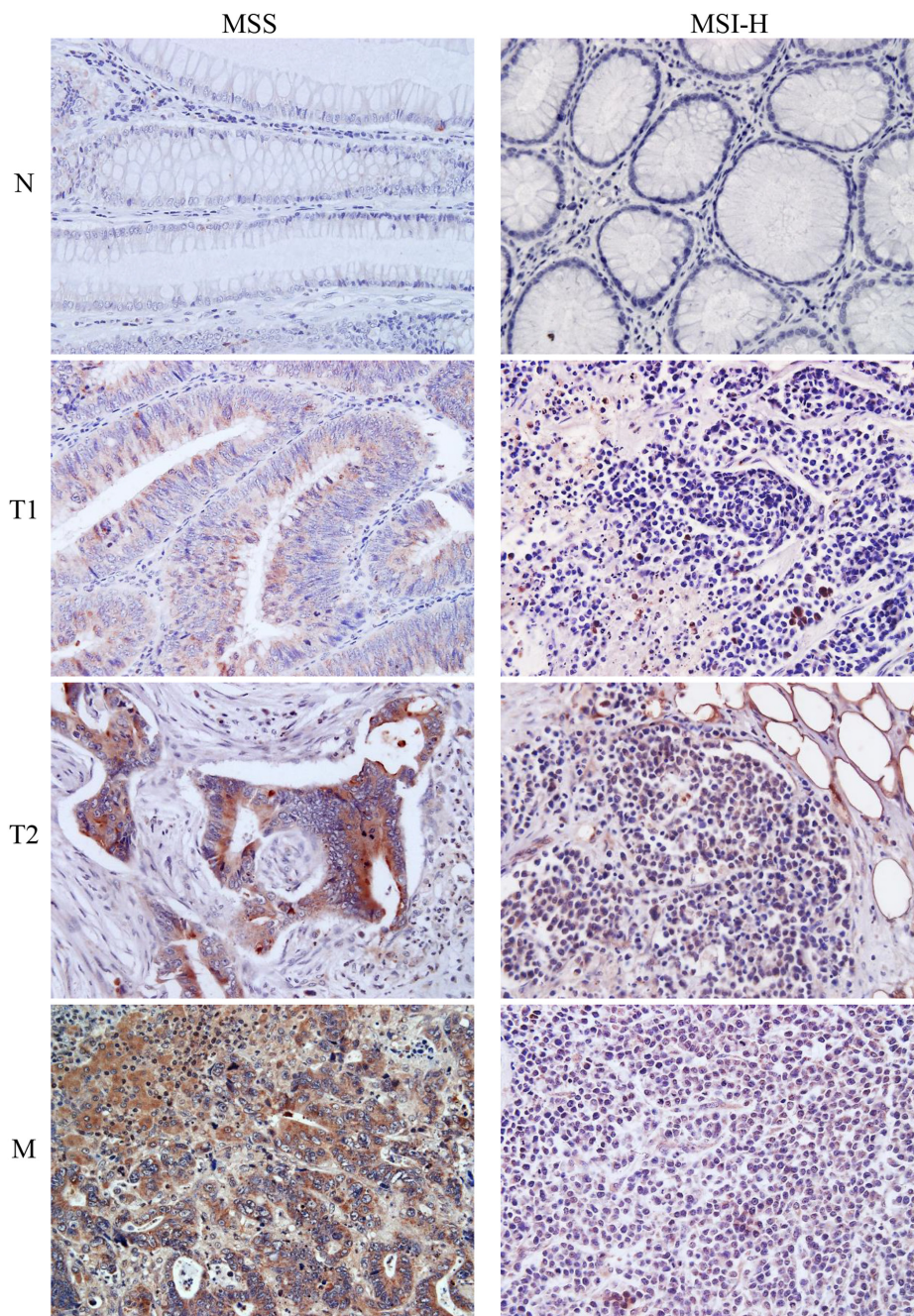


Figure 4. Immunohistochemical staining of epidermal growth factor receptor protein expression. Images show N, T1 and T2 samples from MSS and MSI-H tumours, as well as a liver M sample from an MSS sample and lymph node M sample from an MSI-H tumour sample. N, T1, T2 and M samples were all collected from the same individual (magnification, x400). N, adjacent normal tissue; T1, tumour centre; T2, invasive tumour front; M, metastasis; MSS, microsatellite-stable; MSI-H, microsatellite instability high.

per patient from all samples present in our study. The results showed that *EGFR* mRNA expression levels declined with an

increasing sum of CA-SSR1 alleles which is in line with other papers (22,23), possibly due to changes in DNA secondary

structure. However, this effect was absent at the protein level, an effect seen also by McKay *et al* (41) and Buisine *et al* (28). This could be explained by posttranscriptional regulation via miRNAs (42,43) and possibly other regulatory mechanisms on protein levels like EGFR dimerization, internalization, degradation or recycling (44-46). Interestingly, MSI tumours had lower levels of EGFR on both mRNA and protein levels in comparison to MSS tumours, however, our results show that this is not mediated by *EGFR* CA-SSR1 polymorphism but via some other mechanism that should be further investigated.

To highlight the potential relationship between EGFR overexpression and tumour invasion, we analysed tumour centre and invasive tumour front samples from each patient. Our results showed for the first time that *EGFR* mRNA and protein levels were lower in both the tumour centre and invasive tumour front of MSI-H tumours than in such samples from MSS tumours. This is in accordance with several studies that showed a smaller number of metastatic lymph nodes in MSI-H patients than MSS patients (9,47), and it could also clarify why MSI-H metastatic CRC is rare (48). In addition to increased host immunity (49,50), a decrease in EGFR expression in MSI-H CRC could partially explain why these cancers are less aggressive and have a more promising prognosis. Additionally, in MSS tumours, there was a higher expression of EGFR at the invasive tumour front in comparison to the tumour centre which confirms the putative role of EGFR in tumour invasiveness and the development of metastasis in CRC. This indicates that even though changes in non-coding regions are usually background effects, intron 1 polymorphism could play a role since it is located near the region that has a regulatory function in the *EGFR* gene (51).

Correlation between EGFR expression and colon cancer staging along with histological grade and tumour size is still at issue (8). In our study, we found no association between tumour size, histological grade, or Dukes' stage with a level of EGFR expression either in MSS or in MSI-H tumours. Several studies addressed the possible relationship between EGFR overexpression and tumour stage and/or histological grade. McKay *et al* (41) reported a significant association between the histological grade and EGFR, however, they showed no correlation between EGFR expression and the Dukes' stage. On the other hand, Theodoropoulos *et al* (52) reported an association between advanced tumour stage and high EGFR expression, excluding correlation with tumour grade. Whereas, Del Carmen *et al* (42) showed no correlation with either tumour size, TNM stage or tumour differentiation altogether leading to the conclusion that EGFR remains a controversial prognostic factor. Even though survival analysis for MSS tumours showed no statistical significance, it should be taken in consideration that we had a very small dataset for MSS tumours and a complete lack of survival data for MSI-H tumours. Therefore, this aspect of our research should be expanded in the future.

In conclusion, the present study identified the CA-SSR polymorphism in intron 1 of the *EGFR* gene as a potential new CRC marker that is exclusively altered in MSI-H colorectal tumours. Additionally, we showed higher EGFR expression in MSS than in MSI-H tumours. The highest EGFR expression found in invasive tumour front of MSS tumours suggests more important role of EGFR in MSS tumours progression than in MSI-H tumours where MSI is a dominant molecular

genetic change. Furthermore, instability in *EGFR* CA-SSR1 polymorphism that correlates with decreased EGFR expression suggests that it could serve as an indicator of sporadic MSI-H CRC progression.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Croatian Science Foundation (grant no. HRZZ-IP-2016-06-1430).

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

SK is the PI on the project HRZZ-IP-2016-06-1430 and designed the study. SM, KV, SP and LP performed the experiments and the statistical analyses. AŠ and MP collected the samples and analysed the immunohistological data. SK and SM wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Written informed consent was obtained from all patients included in the present study. The present study was approved by the Ethics Committee of Merkur Clinical Hospital, Zagreb (Zagreb, Croatia) and Medical School, University of Zagreb (Zagreb, Croatia) and was performed in accordance with the ethical standards of the Helsinki Declaration.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM and Wallace MB: Colorectal cancer. *Lancet* 394: 1467-1480, 2019.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68: 394-424, 2018.
3. Cohen S: Isolation and biological effects of an epidermal growth-stimulating protein. *Natl Cancer Inst Monogr* 13: 13-37, 1964.
4. Park JH, Han SW, Oh DY, Im SA, Jeong SY, Park KJ, Kim TY, Bang YJ and Park JG: Analysis of KRAS, BRAF, PTEN, IGF1R, EGFR intron 1 CA status in both primary tumors and paired metastases in determining benefit from cetuximab therapy in colon cancer. *Cancer Chemother Pharmacol* 68: 1045-1055, 2011.
5. Spano JP, Lagorce C, Atlan D, Milano G, Domont J, Benamouzig R, Attar A, Benichou J, Martin A, Morere JF, *et al*: Impact of EGFR expression on colorectal cancer patient prognosis and survival. *Ann Oncol* 16: 102-108, 2005.

6. Rego RL, Foster NR, Smyrk TC, Le M, O'Connell MJ, Sargent DJ, Windschitl H and Sinicrope FA: Prognostic effect of activated EGFR expression in human colon carcinomas: Comparison with EGFR status. *Br J Cancer* 102: 165-172, 2010.
7. Lee WS, Baek JH, Lee JN and Lee WK: Mutations in K-ras and epidermal growth factor receptor expression in Korean patients with stages III and IV colorectal cancer. *Int J Surg Pathol* 19: 145-151, 2011.
8. Ljuslinder I, Melin B, Henriksson ML, Oberg A and Palmqvist R: Increased epidermal growth factor receptor expression at the invasive margin is a negative prognostic factor in colorectal cancer. *Int J Cancer* 128: 2031-2037, 2011.
9. Kang YJ, Jung CK, Choi Y, Lee KY, Kim HJ and Kang WK: Clinicopathologic significances of EGFR expression at invasive front of colorectal cancer. *Korean J Pathol* 44: 16-21, 2010.
10. Ueno H, Mochizuki H, Hashiguchi Y, Shimazaki H, Aida S, Hase K, Matsukuma S, Kanai T, Kurihara H, Ozawa K, *et al*: Risk factors for an adverse outcome in early invasive colorectal carcinoma. *Gastroenterology* 127: 385-394, 2004.
11. Morodomi T, Isomoto H, Shirouzu K, Kakegawa K, Irie K and Morimatsu M: An index for estimating the probability of lymph node metastasis in rectal cancers. Lymph node metastasis and the histopathology of actively invasive regions of cancer. *Cancer* 63: 539-543, 1989.
12. Sohn DK, Chang HJ, Park JW, Choi DH, Han KS, Hong CW, Jung KH, Kim DY, Lim SB, Choi HS and Jeong SY: Histopathological risk factors for lymph node metastasis in submucosal invasive colorectal carcinoma of pedunculated or semipedunculated type. *J Clin Pathol* 60: 912-915, 2007.
13. Kawachi H, Eishi Y, Ueno H, Nemoto T, Fujimori T, Iwashita A, Ajioka Y, Ochiai A, Ishiguro S, Shimoda T, *et al*: A three-tier classification system based on the depth of submucosal invasion and budding/sprouting can improve the treatment strategy for T1 colorectal cancer: A retrospective multicenter study. *Mod Pathol* 28: 872-879, 2015.
14. Shia J, Klimstra DS, Li AR, Qin J, Saltz L, Teruya-Feldstein J, Akram M, Chung KY, Yao D, Paty PB, *et al*: Epidermal growth factor receptor expression and gene amplification in colorectal carcinoma: An immunohistochemical and chromogenic in situ hybridization study. *Mod Pathol* 18: 1350-1356, 2005.
15. Brandt B, Meyer-Staeckling S, Schmidt H, Agelopoulos K and Buerger H: Mechanisms of egfr gene transcription modulation: Relationship to cancer risk and therapy response. *Clin Cancer Res* 12: 7252-7260, 2006.
16. Choi JE, Park SH, Kim KM, Lee WK, Kam S, Cha SI, Kim CH, Kang YM, Kim YC, Han SB, *et al*: Polymorphisms in the epidermal growth factor receptor gene and the risk of primary lung cancer: A case-control study. *BMC Cancer* 7: 199, 2007.
17. Loupakis F, Cremolini C, Fontanini G, Stasi I, Salvatore L and Falcone A: Beyond KRAS: Perspectives on new potential markers of intrinsic and acquired resistance to epidermal growth factor receptor inhibitors in metastatic colorectal cancer. *Ther Adv Med Oncol* 1: 167-181, 2009.
18. Etienne-Grimaldi MC, Pereira S, Magné N, Formento JL, Francoual M, Fontana X, Demard F, Dassonville O, Poissonnet G, Santini J, *et al*: Analysis of the dinucleotide repeat polymorphism in the epidermal growth factor receptor (EGFR) gene in head and neck cancer patients. *Ann Oncol* 16: 934-941, 2005.
19. Frolov A, Liles JS, Kossenkova AV, Tzeng CW, Jhala N, Kulesza P, Varadarajulu S, Eloubeidi M, Heslin MJ and Arnoletti JP: Epidermal growth factor receptor (EGFR) intron 1 polymorphism and clinical outcome in pancreatic adenocarcinoma. *Am J Surg* 200: 398-405, 2010.
20. Kersting C, Agelopoulos K, Schmidt H, Korsching E, August C, Gosheger G, Dirksen U, Juergens H, Winkelmann W, Brandt B, *et al*: Biological importance of a polymorphic CA sequence within intron 1 of the epidermal growth factor receptor gene (EGFR) in high grade central osteosarcomas. *Genes Chromosomes Cancer* 47: 657-664, 2008.
21. Buerger H, Gebhardt F, Schmidt H, Beckmann A, Hutmacher K, Simon R, Lelle R, Boecker W and Brandt B: Length and loss of heterozygosity of an intron 1 polymorphic sequence of egfr is related to cytogenetic alterations and epithelial growth factor receptor expression. *Cancer Res* 60: 854-857, 2000.
22. Amador ML, Oppenheimer D, Perea S, Maitra A, Cusatis G, Iacobuzio-Donahue C, Baker SD, Ashfaq R, Takimoto C, Forastiere A and Hidalgo M: An epidermal growth factor receptor intron 1 polymorphism mediates response to epidermal growth factor receptor inhibitors. *Cancer Res* 64: 9139-9143, 2004.
23. Gebhardt F, Zanker KS and Brandt B: Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. *J Biol Chem* 274: 13176-13180, 1999.
24. Lurje G, Nagashima F, Zhang W, Yang D, Chang HM, Gordon MA, El-Khoueiry A, Husain H, Wilson PM, Ladner RD, *et al*: Polymorphisms in cyclooxygenase-2 and epidermal growth factor receptor are associated with progression-free survival independent of K-ras in metastatic colorectal cancer patients treated with single-agent cetuximab. *Clin Cancer Res* 14: 7884-7895, 2008.
25. Garm Spindler KL, Pallisgaard N, Rasmussen AA, Lindebjerg J, Andersen RF, Crüger D and Jakobsen A: The importance of KRAS mutations and EGFR1A>G polymorphism to the effect of cetuximab and irinotecan in metastatic colorectal cancer. *Ann Oncol* 20: 879-884, 2009.
26. Boland CR and Goel A: Microsatellite instability in colorectal cancer. *Gastroenterology* 138: 2073-2087.e3, 2010.
27. Yuan Z, Shin J, Wilson A, Goel S, Ling YH, Ahmed N, Dopeso H, Jhaver M, Nasser S, Montagna C, *et al*: An A13 repeat within the 3'-untranslated region of epidermal growth factor receptor (EGFR) is frequently mutated in microsatellite instability colon cancers and is associated with increased EGFR expression. *Cancer Res* 69: 7811-7818, 2009.
28. Buisine MP, Wacrenier A, Mariette C, Leteurtre E, Escande F, Aissi S, Ketele A, Leclercq A, Porchet N and Lesuffleur T: Frequent mutations of the CA simple sequence repeat in intron 1 of EGFR in mismatch repair-deficient colorectal cancers. *World J Gastroenterol* 14: 1053-1059, 2008.
29. Spaventi R, Pecur L, Pavelic K, Pavelic ZP, Spaventi S and Stambrook PJ: Human tumour bank in Croatia: A possible model for a small bank as part of the future European tumour bank network. *Eur J Cancer* 30A: 419, 1994.
30. Green MR and Sambrook J: Isolation of high-molecular-weight DNA from mammalian blood using proteinase K and phenol. *Cold Spring Harb Protoc* 2017: pdb prot093492, 2017.
31. Cacev T, Radosevic S, Spaventi R, Pavelic K and Kapitanovic S: NF1 gene loss of heterozygosity and expression analysis in sporadic colon cancer. *Gut* 54: 1129-1135, 2005.
32. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25: 402-408, 2001.
33. Mauri G, Pizzuttilo EG, Amatu A, Bencardino K, Palmeri L, Bonazzina EF, Tosi F, Carlo Stella G, Burrafato G, Scaglione F, *et al*: Retreatment with anti-EGFR monoclonal antibodies in metastatic colorectal cancer: Systematic review of different strategies. *Cancer Treat Rev* 73: 41-53, 2019.
34. Tzeng CW, Frolov A, Frolova N, Jhala NC, Howard JH, Vickers SM, Buchsbaum DJ, Heslin MJ and Arnoletti JP: Pancreatic cancer epidermal growth factor receptor (EGFR) intron 1 polymorphism influences postoperative patient survival and in vitro erlotinib response. *Ann Surg Oncol* 14: 2150-2158, 2007.
35. Shitara M, Sasaki H, Yokota K, Okuda K, Hikosaka Y, Moriyama S, Yano M, Kawaguchi T, Kubo A, Takada M, *et al*: Polymorphisms in intron 1 of the EGFR gene in non-small cell lung cancer patients. *Exp Ther Med* 4: 785-789, 2012.
36. Arifin M, Hiyama K, Tanimoto K, Wiyono WH, Hiyama E and Nishiyama M: EGFR activating aberration occurs independently of other genetic aberrations or telomerase activation in adenocarcinoma of the lung. *Oncol Rep* 17: 1405-1411, 2007.
37. Vashist YK, Trump F, Gebauer F, Kutup A, Güngör C, Kalinin V, Muddasar R, Vettorazzi E, Yekebas EF, Brandt B, *et al*: EGFR intron-1 CA repeat polymorphism is a predictor of relapse and survival in completely resected only surgically treated esophageal cancer. *Target Oncol* 9: 43-52, 2014.
38. Huang SF, Chien HT, Chuang WY, Lai CH, Cheng SD, Liao CT and Wang HM: Epidermal growth factor receptor intron-1 CA repeat polymorphism on protein expression and clinical outcome in Taiwanese oral squamous cell carcinoma. *Sci Rep* 7: 4963, 2017.
39. Han SW, Oh DY, Im SA, Park SR, Lee KW, Song HS, Lee NS, Lee KH, Choi IS, Lee MH, *et al*: Epidermal growth factor receptor intron 1 CA dinucleotide repeat polymorphism and survival of advanced gastric cancer patients treated with cetuximab plus modified FOLFOX6. *Cancer Sci* 101: 793-799, 2010.
40. Zhang W, Weissfeld JL, Romkes M, Land SR, Grandis JR and Siegfried JM: Association of the EGFR intron 1 CA repeat length with lung cancer risk. *Mol Carcinog* 46: 372-380, 2007.

41. McKay JA, Murray LJ, Curran S, Ross VG, Clark C, Murray GI, Cassidy J and McLeod HL: Evaluation of the epidermal growth factor receptor (EGFR) in colorectal tumours and lymph node metastases. *Eur J Cancer* 38: 2258-2264, 2002.
42. Del Carmen S, Corchete LA, Gervas R, Rodríguez A, García M, Alcazar JA, García J, Bengoechea O, Muñoz-Bellvis L, Sayagués JM and Abad M: Prognostic implications of EGFR protein expression in sporadic colorectal tumors: Correlation with copy number status, mRNA levels and miRNA regulation. *Sci Rep* 10: 4662, 2020.
43. Teixeira AL, Gomes M and Medeiros R: EGFR signaling pathway and related-miRNAs in age-related diseases: The example of miR-221 and miR-222. *Front Genet* 3: 286, 2012.
44. Jurisic V, Vukovic V, Obradovic J, Gulyaeva LF, Kushlinskii NE and Djordjevic N: EGFR Polymorphism and survival of NSCLC patients treated with TKIs: A systematic review and meta-analysis. *J Oncol* 2020: 1973241, 2020.
45. Tomas A, Futter CE and Eden ER: EGF receptor trafficking: Consequences for signaling and cancer. *Trends Cell Biol* 24: 26-34, 2014.
46. Seth D, Shaw K, Jazayeri J and Leedman PJ: Complex post-transcriptional regulation of EGF-receptor expression by EGF and TGF-alpha in human prostate cancer cells. *Br J Cancer* 80: 657-669, 1999.
47. Buckowitz A, Knaebel HP, Benner A, Bläker H, Gebert J, Kienle P, von Knebel Doeberitz M and Kloor M: Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. *Br J Cancer* 92: 1746-1753, 2005.
48. Goldstein J, Tran B, Ensor J, Gibbs P, Wong HL, Wong SF, Vilar E, Tie J, Broaddus R, Kopetz S, *et al*: Multicenter retrospective analysis of metastatic colorectal cancer (CRC) with high-level microsatellite instability (MSI-H). *Ann Oncol* 25: 1032-1038, 2014.
49. Xiao Y and Freeman GJ: The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. *Cancer Discov* 5: 16-18, 2015.
50. Kloor M and von Knebel Doeberitz M: The immune biology of microsatellite-unstable cancer. *Trends Cancer* 2: 121-133, 2016.
51. Perucho M: Tumors with microsatellite instability: Many mutations, targets and paradoxes. *Oncogene* 22: 2223-2225, 2003.
52. Theodoropoulos GE, Karafoka E, Papailiou JG, Stamopoulos P, Zambirinis CP, Bramis K, Panoussopoulos SG, Leandros E and Bramis J: P53 and EGFR expression in colorectal cancer: A reappraisal of 'old' tissue markers in patients with long follow-up. *Anticancer Res* 29: 785-791, 2009.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.