International Journal of Innovative Research in Science and Engineering
Volume 08, Issue 05, 2022
www.ijirse.com
ISSN: 2454-9665

Impact of SNPs in microRNAs and their correlation to colorectal cancer susceptibility in north Indian population

Mohammad Taleb Mohsen[a,b], Mausumi Bharadwaj[a] , Seemi Farhat Basir[b*]

[a]*Division of Molecular Genetics & Biochemistry, National Institute of Cancer Prevention & Research (ICMR-NICPR) I-7, Sector-39, Noida, India*

[b]*Department of Biosciences, Jamia Millia Islamia University, New Delhi, India*

***Corresponding Author**

Prof. Seemi Farhat Basir

Department of Biosciences,

Jamia Millia Islamia University, New Delhi, India

Email: sbasir@jmi.ac.in

Abstract

Introduction: Single nucleotide polymorphisms (SNPs), in pre-miRNAs may change the expression of microRNAs, and then their inhibitory role that can help in cancer development. We have studied genotype distribution of *miR-146a*/rs2910164, *miR-196a2*/rs11614913 and *miR-499*/rs3746444 SNPs in association with the risk of colorectal cancer (CRC) development.

Materials and methods: Total of 143 CRC tumor tissues and 143 controls from their adjacent tissues were included. The association between CRC risk and the SNPs genotyping, using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) assay was implemented.

Results: The rs2910164 GC genotype was associated with decreased risk of CRC (GG vs. GC: odd ratio (OR)=3.15; GC vs. CC: OR=0.37,) while G allele and recessive model (GC+CC) were associated with increased CRC risk (G vs. C: OR=1.74; GG vs. GC+CC: OR=2.46). Interestingly, using the CC genotype as a reference, rs11614913 CC vs. TC was associated with an increased risk of CRC (OR=1.69). Moreover, the rs3746444 TT genotype, T allele and T dominant model were associated with increased risk of CRC (TT vs. CC: OR=2.42; T vs. C: OR=1.81; TT+TC vs. CC: OR=2.35) while CC genotype and recessive

model were associated with reduced risk of CRC (CC vs. TC: OR=0.45; TT vs. TC+CC, OR=1.61). Interaction studies for *miR-146a*/*miR-196a2* loci showed that carrying GG/TC is linked with less CRC development. Furthermore, GG/CC and GC/CC loci of *miR-146a*/*miR-499* and CC/CC and TC/CC loci of *miR-196a2*/*miR-499* were linked with decreased CRC risk.

Conclusion: Our findings suggest that an association between rs2910164 and rs3746444 polymorphisms and the etiology of CRC

Keywords*: Colorectal cancer, Single nucleotide polymorphism, rs2910164, rs11614913, rs3746444*

Introduction

Colorectal cancer, CRC, counts for huge burden in the field of human life loss and socioeconomical challenges; as it's considered globally the second most common cancer leading to death with 0.94 million annually, and third most prevalent cancer cases with 1.93 million newly reported cases in 2020 [1]. CRC is more frequent in the developed countries than the developing countries as per the most recent Global Cancer Observatory (GCO) in 2020 [1]. Despite the low frequency of CRC in Indian population; the recent trend of incidence and mortality is noticeably higher than before [2]. CRC is a multifactorial disease, both genetic as well as lifestyle habits are involved in developing it. There are many reported risk factors involved like consumption red meat diet, obesity, smoking and alcohol consumption that have been reported with increased CRC susceptibility [3]. In Indian population, the shift towards intake of processed food, less intake of fruits and vegetables, alcohol consumption, cigarette smoking, low physical activity, moderately increased the risk of CRC which may explain the increased trend in CRC numbers in recent time [4].

The observed improvement in the field of diagnostic biomarkers increases the detection accuracy and overall survival rates for CRC patients. Precise and early diagnosis provide opportunity to start the treatment and monitor through and after the treatment course [5][6]. One of the recent emerged biomarkers in the field of diagnostic biomarkers are microRNAs. MicroRNAs are small well-conserved genetic suppressive regulatory molecules [7], which are produced by RNA polymerase II and Ribonuclease III [8]. MicroRNAs play essential roles as they are involved in most of cellular activities including differentiation, proliferation, cell cycle, apoptosis, host-external genes interaction and so they have crucial role in various cancer types including CRC. Various studies from different populations have shown more microRNAs which are related to CRC prognosis or resistance for medications, considering them either onco-miRs [i.e., miR-196a and mir-196b] [9] or tumor-suppressor [i.e., miR-126 and miR-

133b]. [10] Moreover detailed studies about the role of single nucleotide polymorphisms (SNPs) in miRNAs, have started to take part in diagnostic research as a diagnostic marker also as they affect microRNA's function by affecting binding capability or miRNA stability [11].

miR-146a, miR-196a2 and miR-499 SNPs (rs2910164, rs11614913 and rs3746444) respectively, in pre-miRNAs have been reported to affect their transcripts' stability which ultimately changes carcinogenesis capability due to the regulatory role of miRNAs in targeting various genes [12] [13]. Many clinical reports have shown strong association between rs2910164 and various cancers like; CRC [14], thyroid cancer [15], cervical cancer [12], acute lymphoblastic leukemia [16], breast cancer [17], Prostate cancer [18], hepatocellular carcinoma [19], gastric cancer [20] and more. The elevated expression of miR-196a2 reported with several types of malignancies and so on its rs11614913 SNP, including lung cancer [21], breast cancer [22], cervical cancer [13], gastric cancer [23], esophageal cancer [24], CRC [9], hepatocellular carcinoma, [19] glioma [25] and chronic lymphocytic leukemia [26]. For mir-499 and its rs3746444, limited studies were conducted for cancer patients. It has been reported that rs3746444 is associated with decreased risk of breast cancer [22], and the C allele reportedly associated with cervical cancer [12], glioma, lung cancer, cervical cancer, and hepatocellular carcinoma [25] [27-29].

Therefore, this study has been designed to investigate whether rs2910164, rs11614913 and rs3746444 are associated with CRC susceptibility in Indian population, and further to find out the genotypic variations in correlation with CRC characteristics. To best of our knowledge, this report is the first of its type, in Indian population including targeted SNPs in microRNAs collectively, in tumor tissues and their adjacent tissue controls.**Materials and Methods**

Experimental subjects

The present case-control study included one hundred and forty three patients [105 males and 38 females] with mean age 52.4 ± 15.3 years. histologically confirmed tumor tissue samples and equal number of their adjacent control tissue biopsies from the colon / rectum and rectosigmoid junction from CRC patients attending collaborated hospitals with Jamia Millia Islamia (JMI) university and National Institute of Cancer Prevention and Research (NICPR) in New Delhi, India. Samples from patients matching the inclusion criteria were recruited, where no prior treatment, chemotherapy of radiotherapy, were administered. The provided details of the patients and their demographic, habits, tumor grading along with lymph node invasion and metastasis data were summarized in Table 1. All recorded data was obtained as per the Institutional Review Board and Ethical Committee recommendations [RGCIC ethical approval number:

ISSN: 2454-9665

Res/SCM/31/2018/110 and approval number: 25/7/237/JMI/IEC/2019 from institutional ethical committee in Jamia Millia Islamia] and has been performed in agreement per the ethical standards in the Declaration of Helsinki and its later revisions.

Table 1. Demographic profile for colorectal cancer patients in north Indian population

Characteristics		N=143 (%)
Age (mean ± SD) years		
Male	51.3±15.2	105 (73.4)
Female	55.4±15.5	38 (26.6)
Total	52.4±15.3	143 (100)
Lifestyle factors		
Dwelling		
Urban		85 (59.4)
Rural		58 (40.6)
Smoking		
Non Smoker		91 (63.6)
Smoker		52 (36.4)
Food habit		
Vegetarian		73 (51)
Non vegetarian		70 (49)
Alcohol intake		
No		89(62.2)
Yes		54(37.8)
Tumor location		
Colon		98 (68.5)
Rectum		35 (24.5)
Rectosigmoid Junction		10 (7)
Tumor grade		
1		10 (7)
2		80 (56)
3		39 (27.2)
4		14 (9.8)
Lymph node invasion		
Absent		83(58)
Present		60 (42)
Tumor cell differentiation		
Well & moderately differentiated		97 (67.8)
Poorly & undifferentiated		46 (32.2)

ISSN: 2454-9665

DNA Isolation and Quantification

Genomic DNA was extracted by the standard method with proteinase K digestion, followed by phenol-chloroform extraction [30]. DNA quantity and quality checking was performed by Nanodrop spectrophotometer (*ND1000*), PCR for β-globin amplicon and agarose gel electrophoresis.

International Journal of Innovative Research in Science and Engineering
Volume 08, Issue 05, 2022
www.ijirse.com
ISSN: 2454-9665

Table 2. Characteristics of miR Single Nucleotide Polymorphisms and experimental reaction conditions in the present study

MiRNA	SNP in DNA	Primers		PCR-RFLP			
		Primer sequence	Tm	Annealing Temp.	Amplicon size/ SNP location	Restriction enzyme	Reaction Temp.
Mir-146a	rs2910164 G > C	F: 5'-CATGGGTTGTGTCAGTGTCAGAGCT-3' R: 5'-TGCCTTCTGTCTCCAGTCTTCCAA-3'	69.5°C 68.5°C	62 °C	147 bp/ [+26]	SacI [GAGCT^**C**]	37 °C
Mir-196a2	rs11614913 C > T	F: 5'-CCCCTTCCCTTCTCCTCCAGATA-3' R: 5'-CGAAAACCGACTGATGTAACTCCG-3'	67.6°C 66.3°C	63 °C	149 bp/ [+124]	MspI [C^CG**G**]	37 °C
Mir-499	rs3746444 C > T	F: 5'-CAAAGTCTTCACTTCCCTGCCA-3' R: 5'-GATGTTTAACTCCTCTCCACGTGATC-3'	65.8°C 65.9°C	59 °C	146 bp/ [+119]	BclI [ACTAG^**T**]	55 °C

Determination of microRNA genotypes

The reference sequences of pre-miRNAs and their biological features were obtained from the National Center for Biotechnology Information -NCBI-, and for characterization purpose and to create secondary structures, the Vienna RNA Web Services (http://rna.tbi.univie.ac.at) was used [31]. For SNPs genotypes in miRNAs experimentally identification, Polymerase Chain Reaction-Restriction Fragment Length Polymorphism, PCR-RFLP, method was applied. PCR primers were designed as described earlier by Zhibin Hu et al (2008) [11], thermal cycling conditions were confirmed by gradient PCR. Appropriate restriction enzymes were selected, *SacI*, *MspI* and *BclI* restriction enzymes for miR-146a, miR-196a2 and miR-499 respectively, summarized in Table 2.

Statistical analysis

Statistical analyses were carried out by IBM SPSS Statistics (V. 25, IBM Corp, USA) version 25, and excel 365 version. Fischer's exact test "χ^2" or Chi-square test and analysis of variance (AOVA) test, were used intergroups as needed. Wilks' lambda (Λ) and Levene's tests were implemented to check for the differences between the groups and homogeneity of the data, respectively. The significant p value < 0.05 and 95% Confidence Interval (CI) from the parametric two-tailed Pearson's correlation coefficient were calculated to check the association between different variables accompanied with odd ratios (OR) test. However, not assumed equal variances and Welch tests were applied when homogeneity failed. Additionally, genotyping and allelic frequencies were tested for the accordance of Hardy-Weinberg Equilibrium (HWE).

Results

Subjects Characteristics

There were total of 143 CRC patients, took part in this study, and there were no dropouts, in which tissue samples have been collected along with their cancer free adjacent controls. Subject characteristics are summarized in Table 1. Mean age of all the patients found to be 52.4 ± 15.3 years, with majority of male patients (73.4%), and with higher margin for urban residents (59.4%) over rural inhabitants. Colonic cancer was observed as the major CRC type (98 patients), followed by rectal (24 patient) and 10 rectosigmoidal patients.

The most prevailing risk factors for CRC in our study have reported that 52 smokers (36.4%) out of the total patients, and 54 (37.8%) consume alcohol, while almost half of them are non-vegetarians 73 (51%). With 80 patients (55.9%), stage II was the most common onset diagnosed CRC stage in this study, and 32.2% of the patients were diagnosed with poorly and undifferentiated cancer cells. Advanced cancer stage, where lymphatic nodes invasion, was seen in 83 (58%) of the patients.

Genotyping and alleles frequencies of miRNAs in colorectal cancer samples and controls

miR-146a SNP (rs2910164)

The distribution of genotypes G/C alleles of rs2910164 CRC samples and controls has been shown in Table 3. We have observed that the frequency of (GG) wild homozygous type is higher in tumor samples 74.1% (106/143) as compared to control samples 53.8% (77/143), while the heterozygous genotype (GC) was more than two times frequent in controls 38.5% (55/143) than in tumors 16.8% (24/143). The Variant genotype (CC) was found to be 9.1% (13/143) in tumors as compared to controls 7.7% (11/143). Collectively, p value=0.0002 for these genotypes was statistically significant, which means there is a significant difference in genotype distribution amongst tumor and controls. To find out if there is a dominant or recessive role for the alleles, a combined analysis of G/C dominant mode was done, and (GC+CC) combined genotypes had shown a significant difference in comparison to the wild type (GG) in control group (p value=0.0003; OR [95%CI] =2.46 [1.49-4.04]) but not (GG+GC vs. CC) in which p value=0.6700. Considerably, a significant difference was observed in the codominant mode of inheritance between GG vs. GC tumors with more than 3 folds (p value <0.0001; OR [95%CI] =3.15 [1.80-5.53]) and (p value=0.0368; OR [95%CI] =0.37 [0.14-0.94]) for GC vs. CC. Expectedly from the above findings, allele frequency for G and C alleles was significantly different (p value=0.0066; OR [95%CI] =1.74 [1.16-2.60]). The above findings, suggest that there is association between rs2910164 and CRC decreased risk in Indian population.

miR-196a2 SNP (rs11614913)

The C/T alleles distribution of rs11614913 has also been shown in Table 3. The observed frequency of homozygous genotype (TT) was the same in tumor and control groups, 10.5% (15/143), whereas the heterozygous genotype (TC) was higher in control cases 41.25% (41/143) than in tumors 30.1% (43/143). The variant homozygous genotype (CC) reported in 59.4% (85/143) in cancerous group and 48.25% (69/143) in controls; however, the difference was not

Table 3: Distribution of miRNA Genotypes and Alleles Frequencies in Colorectal cancer patients and controls

Allele model	Genotype	Tumor n (%)	Control n(%)	P value	OR [95% CI]
miR-146a (rs2910164)	GG	106 (74.1)	77 (53.8)		-
	GC	24 (16.8)	55 (38.5)	**0.0002****	
	CC	13 (9.1)	11 (7.7)		
G Dominant allele	GG+GC	130 (90.9)	132 (92.3)	0.67	1 [Reference]
	CC	13 (9.1)	11 (7.7)		0.83 [0.36-1.93]

Recessive allele	GG	106 (74.1)	77 (53.8)	**0.0003****	1 [Reference]
	GC+CC	37 (23.9)	66 (46.2)		2.46 [1.49-4.04]
Allele frequency	G allele	236 (82.5)	209 (73.1)	**0.0066****	1 [Reference]
	C allele	50 (17.5)	77 (26.9)		1.74 [1.16-2.60]
Codominant model	GG vs. GC	106/24	77/55	< **0.0001****	3.15 [1.80-5.53]
	GG vs. CC	106/13	77/11	0.7265	1.16 [0.49-2.74]
	GC vs. CC	24/13	55/11	**0.0368****	0.37 [0.14-0.94]
miR-196 (rs11614913)	CC	85 (59.4)	69 (48.25)		-
	TC	43 (30.1)	59 (41.25)	0.1242	
	TT	15 (10.5)	15 (10.5)		
C Dominant allele	TC+CC	128 (89.5)	128 (89.5)	1	1 [Reference]
	TT	15 (10.5)	15 (10.5)		1.00 [0.47-2.13]
Recessive allele	CC	85 (59.4)	69 (48.3)	0.0577	1 [Reference]
	TT+TC	58 (40.6)	74 (51.7)		1.57 [0.98-2.51]
Allele frequency	C allele	213 (74.5)	197 (68.9)	0.1376	1 [Reference]
	T allele	73 (25.5)	89 (31.1)		1.32 [0.91-1.90]
Codominant model	CC vs. TC	85/43	69/59	**0.0418****	1.69 [1.02-2.80]
	CC vs. TT	85/15	69/15	0.6016	1.23 [0.56-2.69]
	TT vs. TC	15/43	15/59	0.4476	1.37 [0.61-3.10]
miR-499 (rs3746444)	TT	77 (53.8)	60 (42.0)		-
	TC	40 (28.0)	34 (23.7)	**0.0080****	
	CC	26 (18.2)	49 (34.3)		
T Dominant allele	TC+TT	117 (81.8)	94 (65.7)	**0.0023****	1 [Reference]
	CC	26 (18.2)	49 (34.3)		2.35 [1.36-4.06]
Recessive allele	TT	77 (53.8)	60 (42.0)	**0.0442****	1 [Reference]
	CC+TC	66 (46.2)	83 (58.0)		1.61 [1.01-2.57]
Allele frequency	T allele	194 (67.8)	154 (53.8)	**0.0006****	1 [Reference]
	C allele	92 (32.2)	132 (46.2)		1.81 [1.29-2.54]
Codominant model	TT vs. TC	40/77	34/60	0.7643	1.09 [0.62-1.93]
	TT vs. CC	77/26	60/49	**0.0030****	2.42 [1.35-4.33]
	CC vs. TC	26/40	49/34	**0.0180****	0.45 [0.23-0.87]

- OR, odds ratio; CI, confidence interval; P-value, probability from chi-square test; G, Guanine; C, Cytosine; T, Thymine; A, Adenine.
- ** Significant P-values were shown in bold.

significant [p value=0.1242]. Moreover, the OR for C allele to T allele frequency in tumor group was 1.32 times in comparison to control group, while p value=0.1376. Allele carrier analysis has shown identical inheritance of homozygous TT genotype in comparison to TC+CC genotypes. However, C recessive model was close to the minimum cut for significance (p value=0.0577; OR [95%CI] =1.57 [0.98-2.51]) but it did not attain the minimum significance cut off. Not surprisingly, the codominant model did not show any significant difference among TT vs. TC and CC vs. TT genotypes in controls and tumors. By contrast, CC

vs. TC was significantly different between the control and tumor groups (p value=0.0418; OR [95%CI] = 1.69 [1.02-2.80]). One can therefore infer that the TC genotype was less likely to develop CRC in Indian population.

miR-499 SNP (rs3746444)

The genotypes and alleles distribution of rs3746444 in CRC is given in Table 3. It was interesting to observe that there was a significant difference in the distribution of C/T genotypes in all the genotypes, (p value=0.0080). The CC genotype frequency in tumor group was 18.2% (26/143) vs 34.3% (49/143) in controls, and less difference observed for TC genotype, 28.0% vs 23.7% for tumors and controls respectively. The variant TT genotype was also different, 53.8% (77/143) in tumors and 42.0% (60/143) in controls. The dominant and recessive models for C and T alleles had shown significant difference between tumor and controls (p value=0.0023; OR [95%CI] =2.35 [1.36-4.06]) for T dominant allele, and (p value=0.0442; OR [95%CI] =1.61 [1.01-2.57]) for recessive model. On the other hand, the lower frequency of C allele in tumors 32.2% (92/286) and 46.2% (132/286) in controls had shown significant difference in comparison to G allele which is almost two folds more in tumor group as compared to controls (p value=0.0006; OR [95%CI] =1.81 [1.29-2.54]). The codominant models did show significant difference for CC vs. TC and TT vs. CC genotypes with (p value=0.0180; OR [95%CI] =0.45 [0.23-0.87]) and (p value=0.0030; OR [95%CI] =2.42 [1.35-4.33]) respectively. The TT vs. TC model was not significantly different between the groups (p value=0.7643; OR [95%CI] = 1.09 [0.62-1.93]). The above findings suggest that C allele results from rs11614913 may play a protective role against CRC in Indian population.

Combined analysis of SNPs in microRNAs- gene-gene interaction (miR-146a [rs2910164], miR-196a2 [rs11614913] and miR-499 SNP [rs3746444])

Combined miR-SNPs of gene-gene interaction genotypes analysis is shown in Table 4. We have put together the various genotype combinations of tumor group microRNAs (miR-146a/miR-196a2, miR-146a/miR-499 and miR-196a2/miR-499) to find out their interaction and their susceptibility for CRC development risk.

Table 4: Combination Analysis of miR-146, miR-196a2 and miR-499 Polymorphisms in CRC samples

	Genotype	Tumor n (%)	Control n (%)	P value	OR [95% CI]
miR-146a/miR-196a2	GG/CC	67 (46.9)	52 (36.4)	1	1 [Reference]
	GG/TC	29 (20.3)	42 (29.4)	**0.0402****	**1.87 [1.03-3.39]**
	GG/TT	10 (7.0)	12 (8.4)	0.3501	1.55 [0.62-3.86]
	GC/CC	12 (8.4)	9 (6.3)	0.9429	0.97 [0.38-2.47]

	GC/TC	8 (5.6)	12 (8.4)	0.1809	1.93 [0.74-5.07]
	GC/TT	4 (2.8)	3 (2.1)	0.9653	0.97 [0.21-4.51]
	CC/CC	6 (4.2)	8 (5.6)	0.3431	1.72 [0.56-5.26]
	CC/TC	6 (4.2)	5 (3.5)	0.9105	1.07 [0.31-3.71]
	CC/TT	1 (0.7)	0 (0)	0.6061	0.43 [0.02-10.74]
miR-146a/miR-499	GG/TT	59 (41.3)	45 (31.5)	1	1 [Reference]
	GG/TC	27 (18.9)	24 (16.8)	0.6557	1.16 [0.59-2.28]
	GG/CC	20 (14.0)	37 (25.9)	**0.0093****	**2.43 [1.23-4.73]**
	GC/TT	13 (9.1)	12 (8.4)	0.6691	1.21 [0.50-1.90]
	GC/TC	8 (5.6)	3 (2.1)	0.3142	0.49 [0.12-1.96]
	GC/CC	3 (2.1)	9 (6.3)	**0.0489****	**3.93 [1.01-15.37]**
	CC/TT	5 (3.5)	3 (2.1)	0.7511	0.79 [0.18-3.47]
	CC/TC	5 (3.5)	7 (4.9)	0.3258	1.84 [0.55-6.16]
	CC/CC	3 (2.1)	3 (2.1)	0.7471	1.31 [0.25-6.80]
miR-196a2/miR-499	CC/TT	45 (31.5)	32 (22.4)	1	1 [Reference]
	CC/TC	22 (14.5)	22 (34)	0.3696	1.41 [0.67-2.96]
	CC/CC	18 (12.6)	31 (21.7)	**0.0186****	**2.42 [1.16-5.06]**
	TC/TT	22 (15.4)	20 (14.0)	0.5245	1.28 [0.60-2.72]
	TC/TC	15 (10.5)	9 (6.3)	0.7239	0.84 [0.33-2.17]
	TC/CC	6 (4.2)	14 (9.80	**0.0278****	**3.28 [1.14-9.45]**
	TT/TT	10 (7.0)	8 (5.6)	0.8234	1.12 [0.40-3.16]
	TT/TC	3 (2.1)	3 (2.1)	0.6879	1.41 [0.27-7.42]
	TT/CC	2 (1.4)	4 (2.8)	0.2487	2.81 [0.48-16.30]

- OR, odds ratio; CI, confidence interval; P-value, probability from chi-square test; G, Guanine; C, Cytosine; T, Thymine; A, Adenine.
- ** Significant P-values were shown in bold.
- To make the calculation possible in the fields with zero values; 0.5 was added to each value as per Pagano & Gauvreau method

In miR-146a/miR-196a2 combinational analysis and referencing GG/CC wild genotypes, GG/TC genotypes combination are found to be significantly associated with decreased risk against the development of CRC (p value=0.0402; OR [95%CI] =1.87 [1.03-3.39]). CC/TT loci reported zero frequency in control group, so 0.5 was added to each value as per Pagano & Gauvreau [32]; however, it was not significantly different (p value=0.6061). Other allele combinations including GG/TT, GC/TC and CC/CC, indicated towards the decreased risk of CRC by 1.5-2 folds; but their associations were not statistically significant [p-values=0.3501, 0.1809 and 0.607 respectively]. Almost identical scenario was observed among controls

ISSN: 2454-9665

and tumors with ~1 fold change in GG/CC, GC/TT and CC/TC [p-values=0.9429, 0.9653 and 0.9105 respectively].

The interaction of miR-146a with miR-499 SNPs, and referencing GG/TT ancestral wild genotypes, has revealed two significant alleles combinations of that are linked to less CRC risk, GG/CC and GC/CC (p value=0.0093; OR [95%CI] =2.43 [1.23-4.73]) and (p value=0.0489; OR [95%CI] = 3.93 [1.01-15.37]) respectively. There was a slight increase in OR for GG/TC, GC/TT, CC/TT and CC/CC [OR= 1.16, 1.21, 1.84 and 1.31] respectively, and decrease OR for GC/TC and CC/TT [OR= 0.49 and 0.79] respectively; but none of the 6 allele combinations has shown any significant difference between controls and tumors.

Also, analyzing the interaction of miR-196a2 with miR-499 SNPs interaction, referencing CC/TT wild types has exposed two significant values for CC/CC and TC/CC alleles combinations (p value=0.0186; OR [95%CI] =2.42 [1.16-5.06]) and (p value=0.0278; OR [95%CI] = 3.28 [1.14-9.45]) respectively. By contrast, most of the combinations have shown weak association by 1.1-1.4 fold nonsignificant decreased risk of CRC development in CC/TC, TC/TT, TT/TT and TT/TC (OR [95%CI] = 1.41 [0.67-2.96], 1.28 [0.60-2.72], 1.12 [0.40-3.16] and 1.41 [0.27-7.42]) respectively. There was also a nonsignificant decreased CRC risk by 2.81 folds for TT/CC (p value=0.2487; OR [95%CI] =2.81 [0.48-16.30]). Remaining genotype combination TC/TC was found to be non-significantly associated with marginal higher risk of CRC in comparison to CC/TT reference genotype (p value=0.7239; OR [95%CI] =0.84 [0.33-2.17]).

Clinical and lifestyle characteristics association with rs2910164, rs11614913 and rs3746444 genotyping in CRC

As lifestyle and clinical factors have shown correlation with many cancer types, which explains the complexity in cancer prevalence and progression; so, for better understanding of CRC etiology, and the association between lifestyle and clinical factors on rs2910164, rs11614913 and rs3746444 genetic variations, we had tested the association between allelic genotyping pattern in tumor samples with gender, lifestyle habits (smoking, alcohol consumption and food habit) and clinical characteristics in Indian population. Findings and statistical results are provided in supplementary tables 1-3.

rs2910164

The lifestyle habits in association with rs2910164 (G/C) analysis has revealed an increased frequency of CC variant homo-genotype in rs2910164, which is >3 folds in smoker patients than non-smokers (p value=0.0443; OR [95%CI] = 3.39 [1.03-11.13]) and so, the recessive and dominant models have shown two and three folds increase respectively in reference to smoking habit, also p value was close to 0.05, but didn't attain significance (p value=0.0736 for recessive model, and 0.0572 for dominant model). The

association with non-vegetarian food habit has shown significant increase by almost four folds in G dominant allele in comparison to CC homozygous variant allele (p value=0.0462; OR [95%CI] = 3.89 [1.02-14.78]). On the other hand, the three folds increase didn't give significant difference between CC allele and GG reference allele (p value=0.0974). Also, a decrease by ~2.5 folds in reported good GC hetero allele linked with non-vegetarian diet was observed, but it didn't obtain significant difference (p value=0.0698). Alcohol consumption along with clinical factors (gender, lymph node invasiveness, tumor grading, tumor position, tumor cell differentiation and age categorization) have not shown any significant difference among rs2910164 different allele patterns. From the findings, and despite a significant difference between GG vs. CC and GG+GC vs. CC in reference to smoking habit and non-vegetarian food consumption, less correlation was observed between lifestyle habits as well as clinical characteristics with rs2910164 in Indian CRC patients.

rs11614913

The frequency of TT allele and C dominant allele in rs11614913 SNP among CRC female patients have shown an increase by almost 3 folds than male patients, but they didn't attain the minimum significant value (TT vs. CC: p value=0.0702; OR [95%CI] = 2.84 [0.92- 8.82], and TT vs. CC+TC: p value=0.0706; OR [95%CI] = 2.74 [0.92- 8.16]). The correlation between various genotypes with lifestyle habits and clinical characteristics has not shown significant difference between the compared groups (Supplementary table 2). Our findings show similar genotypes distribution among lifestyle and clinical characteristics-based subgroups for rs11614913.

rs3746444

When the comparison was done for miR-499a genotypes regarding rs3746444 (Supplementary table 3) with smokers vs. non-smokers and alcohol consumers vs. non consumers; it didn't show association between those habits and reported genotypes. While non vegetarian food habit has shown two times more TC allele frequency in comparison to vegetarian patients (p value=0.0804; OR [95%CI] = 2.00 [0.92-4.35]). An increased frequency of all rs3746444 genotype models in association with lymph node invasion has been observed, referencing TT allele, [OR: 1.43, 2.04 & 1.65 for TC, CC &TC+CC] respectively, while OR: 1.8 for TC+TT vs. CC. However, the increased frequency didn't attain the minimum significant value for genotype models in comparison to tumor invasion in lymph nodes. Our further analysis based on gender wise, tumor cells differentiation, tumor location, tumor grade and different age grouping didn't show any significant association with rs3746444.

Correlation between lifestyle habits and clinical details with CRC characteristics

ISSN: 2454-9665

The correlation between lifestyle habits and clinical details with CRC characteristics is shown in Table 5. The observed significant habit related to cancer characteristic is of non vegetarian food consumption in correlation with lymph nodal invasion, p value = 0.0248 and OR=0.463 times less, as compared to observed invasiveness among non-vegetarian group with 23.8 % of total cases. Although tumor cell differentiation was poorly diagnosed in one fifth of the total cases, in vegetarian group, and as double OR times than in non-vegetarian, but it was not statistically significant (p value = 0.0771). Other characteristics like tumor location and grading categories also failed to show any significant difference with p value 0.7710 and 0.3697 respectively. Alcohol consumption and smoking have shown less differences between tumor and control groups. In patient who were smokers, p values were found to be 0.2603, 0.7725, 0.4457 and 0.7387 for tumor position, grading, lymph node invasion and cell differentiation respectively. Alcohol consumption didn't differ much with p values 0.8697, 0.4329, 0.4164 and 0.6485 for same mentioned dependent factors respectively. The reported results from males and females were showing more occurance in males, but without any significant difference in the characteristics of CRC cancer (p values = 0.1286, 0.6876, 0.7195 and 0.8449 for tumor position, grading, lymph node invasion and cell differentiation dependent factors respectively) and for

International Journal of Innovative Research in Science and Engineering
Volume 08, Issue 05, 2022
www.ijirse.com
ISSN: 2454-9665

Table 5: Correlation between gender, age category and lifestyle risk factors with CRC tumor characteristics

Characteristics	N	Gender		Age category (Years)				Lifestyle habits		
		Male n(%N)	Female n(%N)	<40 n(%N)	(40-54) n(%N)	(55-69) n(%N)	> 70 n(%N)	Non veg. n(%N)	Smoking n(%N)	Alcohol intake n(%N)
Tumor position										
Colon	98	76 (77.6)	22 (22.4)	22 (22.4)	33 (33.7)	33 (33.7)	10 (10.2)	47 (32.9)	40 (28)	38 (26.6)
Rectum	35	23 (65.7)	12 (34.3)	7 (20)	8 (22.9)	10 (28.6)	10 (28.6)	18 (12.6)	8 (5.6)	12 (8.4)
Rectosigmoid Junction	10	6 (60)	4 (40)	2 (20)	3 (30)	5 (50)	0 (0)	5 (3.5)	4 (2.8)	4 (2.8)
P value		0.1286		0.7058				0.771	0.2603	0.8697
Tumor grade										
1	10	6 (60)	4 (40)	3 (30)	4 (40)	1 (10)	2 (20)	5 (3.5)	6 (4.2)	4 (2.8)
2	80	60 (75)	20 (25)	15 (18.75)	23 (28.75)	30 (37.5)	12 (15)	35 (24.5)	24 (16.8)	28 (19.6)
3	39	32 (82.1)	7 (17.9)	9 (23.1)	13 (33.3)	14 35.9)	3 (7.7)	23 (16.7)	16 (11.2)	15 (10.5)
4	14	7 (50)	7 (50)	4 (28.6)	4 (28.6)	3 (21.4)	3 (21.4)	7 (4.9)	6 (4.2)	7 (4.9)
P value		0.6876		0.9709				0.3697	0.7725	0.4329
Lymph node invasion										
Absent	83	60 (72.3)	23 (27.7)	18 (21.7)	25 (30.1)	31 (37.3)	9 (10.9)	34 (23.8)	28 (19.6)	29 (20.3)
Present	60	45 (75)	15 (25)	13 (21.7)	19 (31.7)	17 (28.3)	11 (18.3)	36 (25.2)	24 (16.8)	25 (17.5)
P value		0.7195		0.5262				**0.0248****	0.4457	0.4164
Tumor cell differentiation										
Well & moderate	96	70 (72.9)	26 (27.1)	19 (19.8)	29 (30.2)	33 (34.4)	15 (15.6)	42 (29.4)	34 (23.8)	35 (24.5)
Poor & undifferentiated	47	35 (74.5)	12 (25.5)	12 (25.5)	15 (31.9)	15 (31.9)	5 (10.7)	28 (19.6)	18 (12.5)	19 (13.3)

ISSN: 2454-9665

P value	0.8449	0.7761	0.0771	0.7387	0.6485

- OR, odds ratio; CI, confidence interval; P-value, probability from chi-square test; G, Guanine; C, Cytosine; T, Thymine; A, Adenine.
- ** Significant P-values were shown in bold.
- To make the calculation possible in the fields with zero values; 0.5 was added to each value as per Pagano & Gauvreau method

ISSN: 2454-9665

same factors, age categorization hasn't shown significant variations (p-values= 0.7058, 9709, 0.5262 and 0.7761). And as observed, that no significant role of lifestyle habits, gender and age categorization as sole factors for CRC development in Indian population, except the strong association with non-vegetarian food habit.

Pre-miRNAs stability

With the help of Vienna RNA secondary structure server and miRNASNP-v3, the secondary structure of pre-miRNAs was constructed and their free energy, ΔG, was reported respectively. The SNP rs2910164 locates in mir-146a-3p, leads to 2.8 units increase in ΔG in the presence of C allele at 60th position in its pre-miRNA form because of C:U pairing mismatch instead of a C:U [33] and so an expressional decrease of leading miR-146a-5p by 1.9-fold as compared to G allele [15]. While the alteration in C allele at the position number 78 in pre-miRNA, rs11614913 SNP, into T allele increases ΔG from -51.20 kCal/mol into -46.60 kCal/mol [33] which tends to decrease its stability and ultimately decreases the expression. The functional mature miR-499 forms have six experimentally confirmed regulated genes only; five genes regulated by miR-499a-5p and one gene by miR-499a-3p [34]. Studies testifying the role of miR-499 in carcinogenesis are few, and so too its rs3746444 SNP at 73rd position in its pre-miRNA. This SNP is located in the seed region of miR-499a-3p; however, T allele has minor change in ΔG in comparison to its reference sequence [33] [Figure 1].

Discussion

Molecular regulation at posttranscriptional level in oncogenes or tumor suppressor genes have been recognized for their role towards cancer development. The present study provides experimental evidence that rs2910164, rs11614913 and rs3746444 in three miRNAs (miR-146a, miR-196a2, miR-499) respectively may play a vital role for CRC susceptibility in north Indian population. Despite many reports about the role of miRNAs-SNPs in cancer, limited reports are available with reference to SNPs in miRNAs and their role in CRC risk [35]. Moreover, the genetic variations reports between tumor tissues and their adjacent controls in CRC are very less [36]. The present study was intended to introduce the first report from tumor tissues and their adjacent controls and so to find out the effect of three SNPs in miRNAs on CRC susceptibility in north Indian population.

International Journal of Innovative Research in Science and Engineering
Volume 08, Issue 05, 2022
www.ijirse.com
ISSN: 2454-9665

Figure 1: Predicted secondary structure of pre-miRNAs and polymorphic pre-miRNAs for *miR-146a, miR-196a2 and miR-499*

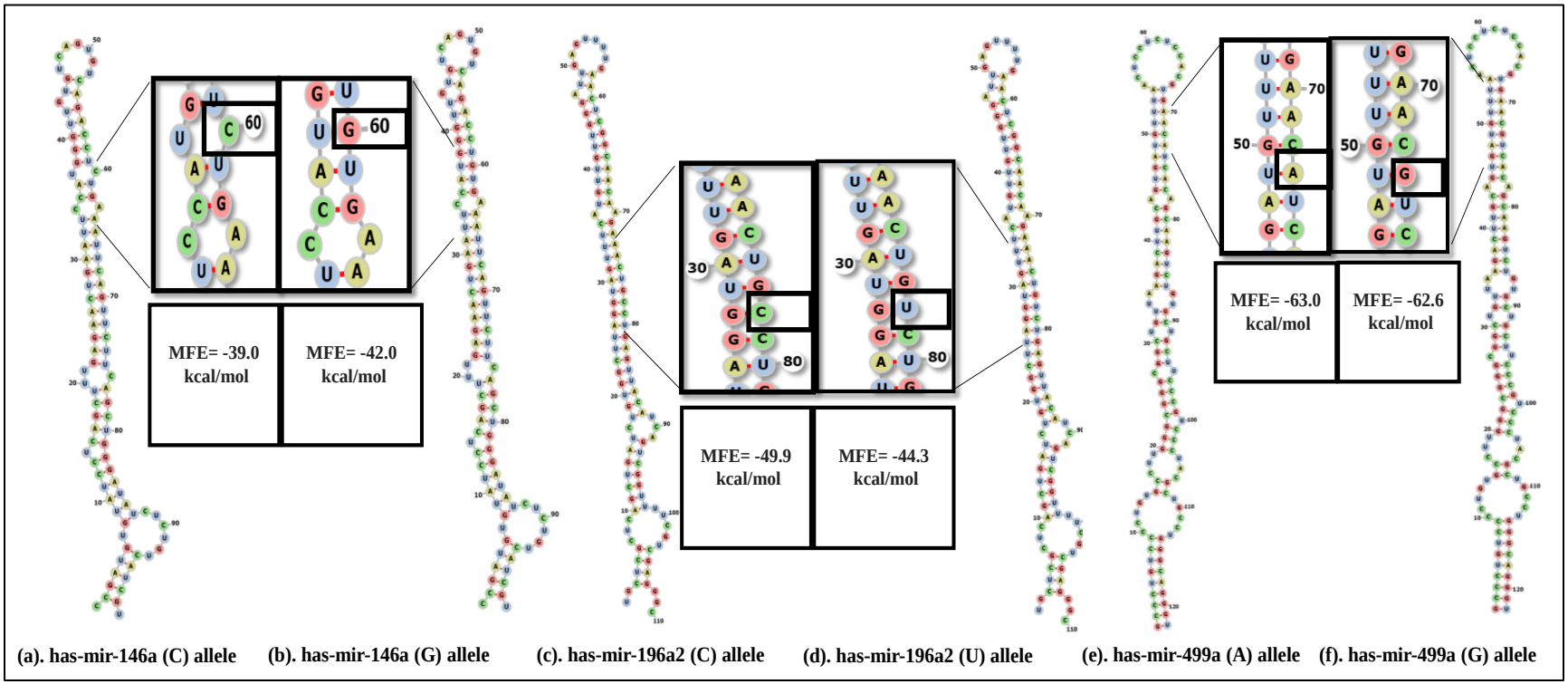

Figure1: The minimum distance between base-pairs as predicted in the centroid secondary structure and minimum free energy (MFE) by the Vienna RNA Web Secondary Structure Server for miRNAs has shown **a**. & **b**. Detailed information for reference *miR-146a* and *miR-146a* rs2910164 polymorphic at 60th nucleotide position secondary structures (C>G), respectively. **c. & d.** Secondary structures for *miR-196a2* and *miR-196a2* rs11614913 polymorphic at 78th nucleotide position (C>U). **e. & f.** Secondary structures for *miR-499* and *miR-499* rs3746444 polymorphic at 73rd nucleotide position (A>G).

ISSN: 2454-9665

A G>C rs2910164 is important since it is located on the complementary, passenger strand, of the functional miR-146a; which could change its stability and so the expression, as we had shown in figure 1 and going in line with what was reported by Jazdzewski et al, (2008) [15] of association between rs2910164 and increased papillary thyroid carcinoma risk, because C allele impacts the expression of miR-146a-5p and its direct targets; IL-1R-associated kinase 1 (IRAKI1), papillary thyroid carcinoma 1 gene (PTC1) and TNFR-associated factor 6 (TRAF6). Strong association between rs2910164 and higher risk for prostate and gastric system cancers development is seen in the presence of homo genotype (GG) [18] [37]. Our finding suggests that the presence of C allele in rs2910164 is significantly associated with protective character against CRC, and more specifically in the presence of GC hetero genotypes amongst Indian population.

In contrast to Jiang et al., 2021, who showed an absence of significant association between rs2910164 and CRC risk in their case control study [38], our study has shown significant protective association between GC hetero genotype and CRC progression to adjacent normal mucosa, in this first report of its type in Indian population.

The increased miR-196a2 expression was reported with multiple cancer types, as well as the SNP rs11614913 which reportedly correlates cancer risk [39]. This study revealed that rs11614913 hetero polymorphic allele was associated with the decreased risk for CRC development. Although tumor tissues with CC allele frequency were higher than controls, but the comparison of CC to the recessive model (TT+TC) couldn't attain the minimum significant value (0.0577) and so the impact of T allele was linked with lower CRC risk but not significantly. Moreover, Wang et al, (2013) showed association between rs11614913 polymorphism and CRC risk but not with specific stage nor location [40].

Other malignancies showed similar protective role for rs11614913 in lung cancer [21], breast cancer [22], cervical cancer [13] and more. The stability of pre-mir196a2 decreases by the change in matching from G:C to G:U in the pre-miRNA form as shown in figure 1 which can play the protective role by decreasing the expression of mir196a2 that has been confirmed with cancer progression and drug resistance [41]. However, lack of association between rs11614913 and colorectal cancer in European population was reported [42] unlike the published data in Asian population [43] and our findings are coherent with that Asian population based findings. Thus, rs11614913 SNP in miR-196a2 could be associated with lower risk for CRC.

In this study, significant association was observed between rs3746444 in miR499 and CRC development risk in Indian population. For all genotype categorization, except the codominant TT vs. TC model, we have reported significant association between T allele and linked genotypes with increased CRC risk. On contrary, C allele and C-linked genotypes have reported lower CRC risk. Despite limited published work

ISSN: 2454-9665

about miR499 and its rs3746444 in CRC, however our first report for Indian population is in line with the findings by Vinci et al. (2013) [44]. By the contrast, Lindor et al. (2017) and Radanova et al. (2022) reported no significant association between the rs3746444 and CRC and further investigation with larger and representative sample size [45] composed of tissue and blood samples as well from different ethnicities. Other studies have shown significant association between rs3746444 and cancer types: breast cancer [46], hepatocellular carcinoma, cervical squamous cell carcinoma, lung cancer and prostate cancer [27]. Our findings suggest that C allele results from rs3746444 may play a protective role against CRC in Indian population.

Synergistic effect of miRNAs and their SNPs has been studied in various cancer types, like oral cancer [47], breast cancer [56] etc. because of growing evidence that gene-gene interaction may impact the link between genes and diseases [48].

This study showed that carrying GG/TC as a result of miR-146a/miR-196a2 loci synergistic effect results with decreased CRC risk. Same effect was observed with GG/CC and GC/CC loci from miR-146a/miR-499 interaction; furthermore, CC/CC and TC/CC loci from miR-196a2/miR-499 interaction were linked with decreased CRC risk and that in agreement with the individual miRNA genotypes. The higher frequency of these loci had ranged between 2-4 folds in controls than tumors. Our findings, therefore, establish the miRNA-miRNA interaction and its association with CRC susceptibility.

Since CRC is multifactorial disease, risk factors can't be overlooked. Smoking, alcohol consumption, non-vegetarian diet etc. are most common risk factors for CRC and other types of cancer as well. The impact of clinical and lifestyle characteristics on genotyping analysis has shown less association with rs2910164, except for smoking habit which showed significant associated with the CC genotype by more than 3 folds in reference to wild type. Also, the dominant model in association with non-vegetarian food has almost reported 4 times higher frequency than in vegetarian group. Other clinical and lifestyle habits didn't show significant association with any of the genetics models. Moreover, the absence of direct relationship in most of lifestyle habits as well as clinical characteristics with CRC; leads to insignificant frequencies mostly for the genotypes, rather a change in ORs only as it was reported by Rawla et al., (2019) in their extensive CRC review [49]. Further analysis [Table S1-3] for the association between lifestyle habits, gender and age categorization haven't shown significant role in CRC development, except the strong association with non-vegetarian food habit in Indian population.

Conclusion

In conclusion, this study is the first of its type in Indian population, incorporating tumor tissue and their adjacent normal tissues to study rs2910164, rs11614913 and rs3746444. Results provide the evidence that

ISSN: 2454-9665

miR-499 polymorphism (rs3746444), individually and in combination with other two SNPs, is associated with a significantly decreased CRC risk. The genotype models for miR-146a (rs2910164) also show significant role for decreasing CRC risk, especially GC genotype. However, miR-196a2 (rs11614913) genetic variant was not associated individually to CRC, but some of its genetic combinations with miR-499 and miR-146a were found to be linked with the lower CRC risk in Indian population. Taken these SNPs individually and synergistically and in correlation with clinical characteristics, may have the prospect for emerging as CRC biomarkers. Extended work with larger studies on miRNAs SNPs in Indian population is needed to fully understand their role in CRC development.

Recommendations and study limitations

Although the results show SNPs linked to CRC risk because of various genotypes and loci; we recommend comprehensive large studies including different ethnicities and to include blood as well as tissue samples. Limitation of the present study is the limited sample size and recruited samples were collected from New Delhi hospitals only.

Acknowledgements: The authors acknowledge infrastructure support of ICMR-NICPR, Noida. We also gratefully acknowledge patients, their relatives, collaborated hospitals and ethical committees from RGCIC and Jamia Millia Islamia for their support and co-operation.

Conflict of interest: Authors report no potential conflicts of interest.

International Journal of Innovative Research in Science and Engineering
Volume 08, Issue 05, 2022
www.ijirse.com
ISSN: 2454-9665

Supplementary table 1: Correlation between clinical characteristics and lifestyle risk factors with *miR-146a* SNP (rs2910164) in CRC

Characteristics	N=143	mir 146a Tumor			Recessive model	Dominant model
		GG [n (%N)]	GC [n (%N)]	CC [n (%N)]	GC+CC vs. GG [n (%N)]	GG+GC* vs. CC [n (%N)]
Gender						
Male	105	80 (55.9)	15 (10.5)	10 (7.0)	25 (17.5)	95 (66.4)
Female	38	26 (18.2)	9 (6.3)	3 (2.1)	12 (8.4)	35 (24.5)
P value		1	0.1999	0.9084	0.3487	0.7667
OR [95% CI]		1 [Reference]	1.85 [0.72-4.71]	0.92 [0.24-3.61]	1.48 [0.65-3.35]	0.81 [0.21-3.13]
Lifestyle factors						
Smoking						
Non smoker	91	72 (50.3)	14 (9.8)	5 (3.5)	19 (13.3)	86 (60.1)
Smoker	52	34 (23.8)	10 (7.0)	8 (5.6)	18 (12.6)	44 (30.8)
P value		1	0.3718	**0.0443****	0.0736	0.0572
OR [95% CI]		1 [Reference]	1.51 [0.61-3.75]	**3.39 [1.03-11.13]**	2.01 [0.94-4.30]	3.13 [0.97-10.13]
Food habit						
Vegetarian	73	53 (37.1)	17 (11.9)	3 (2.1)	20 (14)	70 (49.0)
Non vegetarian	70	53 (37.1)	7 (4.9)	10 (7.0)	17 (11.9)	60 (42.0)
P value		1	0.0698	0.0794	0.6712	0.0462
OR [95% CI]		1 [Reference]	0.41 [0.16-1.07]	3.33 [0.87-12.80]	0.85 [0.40-1.80]	3.89 [1.02-14.78]
Alcohol intake						
No	89	70 (49.0)	13 (9.1)	6 (4.2)	19 (13.3)	83 (58.1)
Yes	54	36 (25.2)	11 (7.7)	7 (4.9)	18 (12.6)	46 (32.9)
P value		1	0.2771	0.1671	0.115	0.2039
OR [95% CI]		1 [Reference]	1.64 [0.67-4.04]	2.27 [0.71-7.25]	1.84 [0.86-3.94]	2.10 [0.67-6.64]
Lymph node invasion						
Absent	83	64 (44.7)	12 (8.4)	7 (4.9)	19 (13.3)	76 (53.1)
Present	60	42 (29.4)	12 (8.4)	6 (4.2)	18 (12.6)	54 (37.8)

P value		1	0.3535	0.6512	0.3393	0.7481
OR [95% CI]		1 [Reference]	1.52 [0.63-3.71]	1.31 [0.41-4.16]	1.44 [0.68-3.07]	1.21 [0.38-3.79]
Tumor cell differentiation						
Well & moderately differentiated	97	71 (49.7)	14 (9.8)	11 (7.7)	25 (17.5)	85 (59.4)
Poorly & undifferentiated	46	35 (24.5)	10 (7.0)	2 (1.4)	12 (8.4)	45 (31.5)
P value		1	0.4228	0.2102	0.9479	0.1764
OR [95% CI]		1 [Reference]	1.45 [0.58-3.59]	0.37 [0.08-1.75]	0.97 [0.44-2.16]	0.34 [0.07-1.62]
Tumor location						
Colon	98	73 (51.0)	14 (9.8)	11 (7.7)	25 (17.5)	87 (60.8)
Rectum	35	24 (16.8)	9 (6.3)	2 (1.4)	11 (7.7)	33 (23.1)
Rectosigmoid Junction	10	9 (6.3)	1 (0.7)	0 (0)	1 (0.7)	10 (7.0)
P value			0.0843		0.1372	0.3689
Tumor grade						
Stage 1	10	7 (4.9)	2 (1.4)	1 (0.7)	3 (2.1)	9 (6.3)
Stage 2	80	60 (42.0)	11 (7.7)	9 (6.3)	20 (14)	71 (49.7)
Stage 3	39	30 (21.0)	8 (5.6)	1 (0.7)	9 (6.3)	38 (26.6)
Stage 4	14	9 (6.3)	3 (2.1)	2 (1.4)	5 (3.5)	12 (8.4)
P value			0.6449		0.6743	0.2252
Age category						
<40 Years	31	23 (16.1)	6 (4.2)	2 (1.4)	8 (5.6)	29 (20.3)
(40-54) Years	44	32 (22.4)	6 (4.2)	6 (4.2)	12 (8.4)	38 (26.6)
(55-69) Years	48	38 (26.6)	7 (4.9)	3 (2.1)	10 (7.0)	45 (31.5)
> 70 Years	20	13 (9.1)	5 (3.5)	2 (1.4)	7 (4.9)	18 (12.6)
P value			0.6537		0.5257	0.6113

- OR, odds ratio; CI, confidence interval; P-value, probability from chi-square test; G, Guanine; C, Cytosine; T, Thymine; A, Adenine.
- ** Significant P-value was shown in bold.
- To make the calculation possible in the fields with zero values; 0.5 was added to each value as per Pagano & Gauvreau metho

Supplementary table 2: Correlation between clinical characteristics and lifestyle risk factors with *miR-196a2* SNP (rs11614913) in CRC

International Journal of Innovative Research in Science and Engineering
Volume 08, Issue 05, 2022
www.ijirse.com
ISSN: 2454-9665

Characteristics	N=143	mir 196a2 Tumor			Recessive model	Dominant model
		CC [n (%N)]	TC [n (%N)]	TT [n (%N)]	TT+TC vs.CC [n (%N)]	CC+TC* vs. TT [n (%N)]
Gender						
Male	105	65 (45.5)	32 (22.4)	8 (5.6)	40 (28.0)	97 (67.8)
Female	38	20 (14.0)	11 (7.7)	7 (4.9)	18 (12.6)	31 (21.7)
P value		1	0.798	0.0702	0.3197	0.0706
OR [95% CI]		1 [Reference]	1.12 [0.48-2.61]	2.84 [0.92-8.82]	1.46 [0.69-3.09]	2.74 [0.92-8.16]
Lifestyle factors						
Smoking						
Non smoker	91	27 (18.9)	56 (39.2)	8 (5.6)	64 (44.8)	83 (58.0)
Smoker	52	16 (11.2)	29 (20.3)	7 (4.9)	36 (25.2)	45 (31.5)
P value		1	0.7294	0.5203	0.8904	0.3839
OR [95% CI]		1 [Reference]	0.87 [0.41-1.88]	1.48 [0.45-4.84]	0.95 [0.45-1.99]	1.61 [0.55-4.74]
Food habit						
Vegetarian	73	43 (30.1)	20 (14.0)	10 (7.0)	30 (21.0)	63 (44.4)
Non vegetarian	70	42 (29.4)	23 (16.1)	5 (3.5)	24 (16.8)	65 (45.4)
P value		1	0.6631	0.2557	0.5678	0.2082
OR [95% CI]		1 [Reference]	1.18 [0.56-2.45]	0.51 [0.16-1.62]	0.82 [0.41-1.62]	0.48 [0.16-1.50]
Alcohol intake						
No	89	55 (38.5)	23 (16.1)	11 (7.7)	34 (23.8)	78 (45.1)
Yes	54	30 (21.0)	20 (14.0)	4 (2.8)	24 (16.8)	50 (35.0)
P value		1	0.2207	0.5175	0.4615	0.3537
OR [95% CI]		1 [Reference]	1.59 [0.76-3.36]	0.67 [0.19-2.28]	1.29 [0.56-2.57]	0.57 [0.17-1.88]
Lymph node invasion						
Non invasive	83	51 (35.7)	24 (16.8)	8 (5.6)	32 (22.4)	75 (52.4)
Invasive	60	34 (23.8)	19 (13.3)	7 (4.9)	26 (18.2)	53 (37.1)
P value		1	0.6499	0.629	0.5659	0.6965

OR [95% CI]		1 [Reference]	1.19 [0.56-2.49]	1.31 [0.43-3.96]	1.22 [0.62-2.39]	1.24 [0.42-3.62]
Tumor cell differentiation						
Well & moderately differentiated	97	54 (37.8)	32 (22.4)	10 (7.0)	42 (29.4)	86 (60.1)
Poorly & undifferentiated	46	31 (21.7)	11 (7.7)	5 (3.5)	16 (11.2)	42 (29.4)
P value		1	0.2175	0.8156	0.268	0.9676
OR		1 [Reference]	0.60 [0.26-1.35]	0.87 [0.27-2.78]	0.66 [0.32-1.371]	1.02 [0.32-3.19]
Tumor location						
Colon	98	56 (39.2)	34 (23.8)	8 (5.6)	42 (29.4)	90 (62.9)
Rectum	35	21 (14.7)	8 (5.6)	6 (4.2)	14 (9.8)	29 (20.3)
Rectosigmoid Junction	10	8 (5.6)	1 (0.7)	1 (0.7)	2 (1.4)	9 (6.3)
P value			0.5418		0.3035	0.4655
Tumor grade						
Stage 1	10	5 (3.5)	3 (2.1)	2 (1.4)	8 (5.6)	5 (3.5)
Stage 2	80	48 (33.6)	25 (17.5)	7 (4.9)	73 (51.0)	32 (22.4)
Stage 3	39	24 (16.8)	10 (7.0)	5 (3.5)	34 (23.8)	15 (10.5)
Stage 4	14	8 (5.6)	5 (3.5)	1 (0.7)	13 (9.1)	6 (4.2)
P value			0.8355		0.9252	0.6632
Age category						
<40 Years	31	19 (13.3)	9 (6.3)	3 (2.1)	28 (19.6)	12 (8.4)
(40-54) Years	44	29 (20.3)	12 (8.4)	3 (2.1)	41 (28.7)	15 (10.5)
(55-69) Years	48	28 (19.6)	15 (10.5)	5 (3.5)	43 (30.1)	20 (14.0)
> 70 Years	20	9 (6.3)	7 (4.9)	4 (2.8)	16 (11.2)	11 (7.7)
P value			0.3213		0.4713	0.6178

- OR, odds ratio; CI, confidence interval; P-value, probability from chi-square test; G, Guanine; C, Cytosine; T, Thymine; A, Adenine.
- To make the calculation possible in the fields with zero values; 0.5 was added to each value as per Pagano & Gauvreau method

Supplementary table 3: Correlation between clinical characteristics and lifestyle risk factors with *miR-499* SNP (rs3746444) genotypes in CRC

International Journal of Innovative Research in Science and Engineering
Volume 08, Issue 05, 2022
www.ijirse.com
ISSN: 2454-9665

Characteristics	N=143	mir 499 Tumor			Recessive model	Dominant model
		TT [n (%N)]	TC [n (%N)]	CC [n (%N)]	CC+TC vs.TT [n (%N)]	TC+TT* vs. CC [n (%N)]
Gender						
Male	105	57 (36.6)	28 (19.6)	20 (14.0)	48 (33.6)	85 (59.4)
Female	38	20 (14.0)	12 (8.4)	6 (4.2)	18 (12.6)	32 (22.4)
P value		1	0.6433	0.7689	0.8609	0.6559
OR [95% CI]		1 [Reference]	1.22 [0.52-2.85]	0.85 [0.30-2.43]	0.94 [0.51-2.25]	0.80 [0.29-2.16]
Lifestyle factors						
Smoking						
Non smoker	91	46 (32.2)	26 (18.2)	19 (13.3)	45 (31.5)	72 (50.3)
Smoker	52	31 (21.7)	14 (9.8)	7 (4.9)	21 (14.7)	45 (31.5)
P value		1	0.5794	0.2267	0.2964	0.2721
OR [95% CI]		1 [Reference]	0.80 [0.36-1.77]	0.55 [0.20-1.45]	0.69 [0.34-1.38]	0.59 [0.23-1.51]
Food habit						
Vegetarian	73	44 (30.8)	16 (11.2)	13 (9.1)	29 (20.3)	60 (42.0)
Non vegetarian	70	33 (23.1)	24 (16.8)	13 (9.1)	37 (25.9)	57 (40.0)
P value		1	0.0804	0.5271	0.1164	0.9058
OR [95% CI]		1 [Reference]	2.00 [0.92-4.35]	1.33 [0.55-3.25]	1.70 [0.88-3.30]	1.05 [0.45-2.46]
Alcohol intake						
No	89	48 (33.6)	24 (16.8)	17 (11.9)	41 (28.7)	72 (50.3)
Yes	54	29 (20.3)	16 (11.2)	9 (6.3)	25 (17.5)	45 (31.5)
P value		1	0.8053	0.7808	0.9788	0.7146
OR [95% CI]		1 [Reference]	1.10 [0.50-2.41]	0.88 [0.35-2.22]	1.01 [0.51-1.99]	0.85 [0.35-2.06]

Lymph node invasion						
Non invasive	83	49 (34.3)	22 (15.4)	12 (8.4)	34(23.8)	71 (49.6)
Invasive	60	28 (19.6)	18 (12.6)	14 (9.8)	32 (22.4)	46 (25.2)
P value		1	0.3652	0.1201	0.1442	0.1779
OR [95% CI]		1 [Reference]	1.43 [0.66-3.11]	2.04 [0.83-5.02]	1.65 [0.84-3.21]	1.80 [0.76-4.24]
Tumor cell differentiation						
Well & moderately differentiated	97	53 (37.1)	27 (18.9)	16 (11.2)	43 (30.1)	80 (55.9)
Poorly & undifferentiated	46	24 (16.8)	13 (9.1)	10 (7.0)	23 (16.1)	37 (25.9)
P value		1	0.8832	0.495	0.6406	0.5029
OR		1 [Reference]	1.06 [0.47-2.41]	1.38 [0.55-3.48]	1.18 [0.59-2.38]	1.35 [0.56-3.26]
Tumor location						
Colon	98	53 (37.1)	25 (17.5)	20 (14.0)	45 (31.5)	78 (54.5)
Rectum	35	18 (12.6)	12 (8.4)	5 (3.5)	17 (11.9)	30 (21.0)
Rectosigmoid Junction	10	6 (4.2)	3 (2.1)	1 (0.7)	4 (2.8)	9 (6.3)
P value			0.812		0.8904	0.5728
Tumor grade						
Stage 1	10	4 (2.8)	4 (2.8)	2 (1.4)	6 (4.2)	8 (5.6)
Stage 2	80	45 (31.5)	21 (14.7)	14 (9.8)	35 (24.5)	66 (46.1)
Stage 3	39	18 (12.6)	14 (9.8)	7 (4.9)	21 (14.7)	32 (22.4)
Stage 4	14	10 (7.0)	1 (0.7)	3 (2.1)	4 (2.8)	11 (7.7)
P value			0.7136		0.327	0.986
Age category						
<40 Years	31	18 (12.6)	9 (6.3)	4 (2.8)	13 (9.1)	27 (18.9)
(40-54) Years	44	20 (14.0)	14 (9.8)	10 (7.0)	24 (16.8)	34 (23.8)
(55-69) Years	48	27 (18.9)	11 (7.7)	10 (7.0)	21 (14.7)	38 (26.6)
> 70 Years	20	12 (8.4)	6 (4.2)	2 (1.4)	8 (5.6)	18 (12.6)

ISSN: 2454-9665

P value	0.4997	0.6039	0.4542

- OR, odds ratio; CI, confidence interval; P-value, probability from chi-square test; G, Guanine; C, Cytosine; T, Thymine; A, Adenine.
- To make the calculation possible in the fields with zero values; 0.5 was added to each value as per Pagano & Gauvreau method

References:

[1]. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, *71*(3), 209-249

[2]. Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International Journal of Cancer*, *149*(4), 778-789

[3]. Johnson, C. M., Wei, C., Ensor, J. E., Smolenski, D. J., Amos, C. I., Levin, B., & Berry, D. A. (2013). Meta-analyses of colorectal cancer risk factors. *Cancer causes & control*, *24*(6), 1207-1222.

[4]. Sinha, R., Doval, D. C., Hussain, S., Kumar, K., Singh, S., Basir, S. F., & Bharadwaj, M. (2015). Lifestyle and sporadic colorectal cancer in India. *Asian Pacific Journal of Cancer Prevention*, *16*(17), 7683-7688.

[5]. Patel, J. N., Fong, M. K., & Jagosky, M. (2019). Colorectal cancer biomarkers in the era of personalized medicine. *Journal of personalized medicine*, *9*(1), 3

[6]. Pang, S. W., Awi, N. J., Armon, S., Lim, W. W. D., Low, J. S. H., Peh, K. B., & Teow, S. Y. (2020). Current update of laboratory molecular diagnostics advancement in management of colorectal cancer (CRC). *Diagnostics*, *10*(1), 9

[7]. Engels, B. M., & Hutvagner, G. (2006). Principles and effects of microRNA-mediated post-transcriptional gene regulation. *Oncogene*, *25*(46), 6163-6169.

[8]. Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., & Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal*, *23*(20), 4051-4060

[9]. Ge, J., Chen, Z., Li, R., Lu, T., & Xiao, G. (2014). Upregulation of microRNA-196a and microRNA-196b cooperatively correlate with aggressive progression and unfavorable prognosis in patients with colorectal cancer. *Cancer cell international*, *14*(1), 1-8

[10]. Masuda, T., Hayashi, N., Kuroda, Y., Ito, S., Eguchi, H., & Mimori, K. (2017). MicroRNAs as biomarkers in colorectal cancer. *Cancers*, *9*(9), 124

[11]. Moszyńska, A., Gebert, M., Collawn, J. F., & Bartoszewski, R. (2017). SNPs in microRNA target sites and their potential role in human disease. *Open biology*, *7*(4), 170019

[12]. Thakur, N., Singhal, P., Mehrotra, R., & Bharadwaj, M. (2019). Impacts of single nucleotide polymorphisms in three microRNAs (miR-146a, miR-196a2 and miR-499) on the susceptibility to cervical cancer among Indian women. *Bioscience reports*, *39*(4).

[13]. Liu, J., Dong, P., Zhou, L., & Wang, S. (2021). The Association between Five Genetic Variants in MicroRNAs (rs2910164, rs11614913, rs3746444, rs11134527, and rs531564) and Cervical Cancer Risk: A Meta-Analysis. *BioMed research international*, *202*.

[14]. Lu, D., Yao, Q., Zhan, C., Le-Meng, Z., Liu, H., Cai, Y., & Zhang, S. (2017). MicroRNA-146a promote cell migration and invasion in human colorectal cancer via carboxypeptidase M/src-FAK pathway. *Oncotarget*, *8*(14), 22674.

[15]. Jazdzewski, K., Murray, E. L., Franssila, K., Jarzab, B., Schoenberg, D. R., & de la Chapelle, A. (2008). Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. *Proceedings of the National Academy of Sciences*, *105*(20), 7269-7274.

[16]. Zou, D., Yin, J., Ye, Z., Zeng, Q., Tian, C., Wang, Y., & Chen, R. (2020). Association Between the miR-146a Rs2910164 Polymorphism and Childhood Acute Lymphoblastic Leukemia Susceptibility in an Asian Population. *Frontiers in genetics*, *11*, 886.

[17]. Yi, D., Wang, R., Shi, X., Xu, L., Yilihamu, Y. E., & Sang, J. (2020). METTL14 promotes the migration and invasion of breast cancer cells by modulating N6-methyladenosine and hsa-miR-146a-5p expression. *Oncology reports*, *43*(5), 1375-1386

[18]. Liu, Y., Gao, S., Du, Q., & Shao, M. (2019). miR-146a and miR-152 in prostate cancer and clinicopathological parameters. *J Buon*, *24*, 1692-1699

[19]. Zheng, L., Zhuang, C., Zhao, J., & Ming, L. (2017). Functional miR-146a, miR-149, miR-196a2 and miR-499 polymorphisms and the susceptibility to hepatocellular carcinoma: an updated meta-analysis. *Clinics and research in hepatology and gastroenterology*, *41*(6), 664-676.

[20]. Shomali, N., Shirafkan, N., Duijf, P. H., Ghasabi, M., Babaloo, Z., Yousefi, M., & Baradaran, B. (2019). Downregulation of miR-146a promotes cell migration in Helicobacter pylori–negative gastric cancer. *Journal of cellular biochemistry*, *120*(6), 9495-9505.

[21]. Hu, Z., Chen, J., Tian, T., Zhou, X., Gu, H., Xu, L., & Shen, H. (2008). Genetic variants of miRNA sequences and non–small cell lung cancer survival. *The Journal of clinical investigation*, *118*(7), 2600-2608

[22]. Morales, S., De Mayo, T., Gulppi, F. A., Gonzalez-Hormazabal, P., Carrasco, V., Reyes, J. M., & Jara, L. (2018). Genetic Variants in pre-miR-146a, pre-miR-499, pre-miR-125a, pre-miR-605, and pri-miR-182 Are Associated with Breast Cancer Susceptibility in a South American Population. *Genes*, *9*(9), 427.
[23]. Poltronieri-Oliveira, A. B., Madeira, F. F., Nunes, D. B. S. M., Rodrigues, G. H., Lopes, B. C., Manoel-Caetano, F. S., & Silva, A. E. (2017). Polymorphisms of miR-196a2 (rs11614913) and miR-605 (rs2043556) confer susceptibility to gastric cancer. *Gene Reports*, *7*, 154-163.
[24]. Guo, X., Zhao, L., Shen, Y., Shao, Y., Wei, W., & Liu, F. (2019). Polymorphism of miRNA and esophageal cancer risk: an updated systemic review and meta-analysis. *OncoTargets and therapy*, *12*, 3565.
[25]. Yang, S., Zheng, Y., Zhou, L., Jin, J., Deng, Y., Yao, J., & Dai, Z. (2020). miR-499 rs3746444 and miR-196a-2 rs11614913 Are Associated with the Risk of Glioma, but Not the Prognosis. *Molecular Therapy-Nucleic Acids*, *22*, 340-351
[26]. Deghady, A., Elwafa, R. A., Lsorady, M. A. E., & Elbana, D. (2019). MicroRNA-196a2 single nucleotide polymorphism rs11614913 in Egyptian patients with chronic lymphocytic leukemia. *Hematol Transfus Int J*, *7*(1), 17-20
[27]. Yang, X., Li, X., & Zhou, B. (2018). A meta-analysis of miR-499 rs3746444 polymorphism for cancer risk of different systems: evidence from 65 case-control studies. *Frontiers in physiology*, *9*, 737
[28]. Yan, Z., Zhou, Z., Li, C., Yang, X., Yang, L., Dai, S., & Yao, Y. (2019). Polymorphisms in miRNA genes play roles in the initiation and development of cervical cancer. *Journal of Cancer*, *10*(20), 4747
[29]. Nouri, R., & Ghorbian, S. (2019). Association of single nucleotide polymorphism in hsa-miR-499 and hsa-miR-196a2 with the risk of prostate cancer. *International urology and nephrology*, *51*(5), 811-816.
[30]. Sambrook J, Fritsh EF & Maniatis T. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1989. 9.16-9.19
[31]. Hofacker, I. L. (2003). Vienna RNA secondary structure server. *Nucleic acids research*, *31*(13), 3429-3431
[32]. Pagano, M., & Gauvreau, K. (2018). *Principles of biostatistics*. CRC Press
[33]. Ghun-jie Liu, Xin Fu, Mengxuan Xia, Qiong Zhang, Zhifeng Gu & An-yuan Guo. (2020) miRNASNP-v3: a comprehensive database for SNPs and disease-related variations in miRNAs and miRNA targets. *Nucleic Acids Research*.
[34]. Kern, F., Aparicio-Puerta, E., Li, Y., Fehlmann, T., Kehl, T., Wagner, V., & Keller, A. (2021). miRTargetLink 2.0—interactive miRNA target gene and target pathway networks. *Nucleic Acids Research*, *49*(W1), W409-W416
[35]. To, K. K., Tong, C. W., Wu, M., & Cho, W. C. (2018). MicroRNAs in the prognosis and therapy of colorectal cancer: From bench to bedside. *World journal of gastroenterology*, *24*(27), 2949.
[36]. Xing, J., Wan, S., Zhou, F., Qu, F., Li, B., Myers, R. E., & Yang, H. (2012). Genetic polymorphisms in pre-microRNA genes as prognostic markers of colorectal cancer. *Cancer Epidemiology and Prevention Biomarkers*, *21*(1), 217-227
[37]. Park, R., Lopes, L., & Saeed, A. (2020). Association between the functional miR-146a SNP rs2910164 and risk of digestive system cancer: updated meta-analysis. *Anticancer research*, *40*(3), 1495-1502
[38]. Jiang, J., Zhang, S., Tang, W., & Qiu, Z. (2021). Lack of association between miR-146a rs2910164 C/G locus and colorectal cancer: from a case–control study to a meta-analysis. *Bioscience Reports*, *41*(1), BSR20191729
[39]. Choupani, J., Nariman-Saleh-Fam, Z., Saadatian, Z., Ouladsahebmadarek, E., Masotti, A., & Bastami, M. (2019). Association of mir-196a-2 rs11614913 and mir-149 rs2292832 polymorphisms with risk of cancer: an updated meta-analysis. *Frontiers in genetics*, *10*, 186.
[40]. Wang, N., Li, Y., Zhu, L. J., Zhou, R. M., Jin, W., Guo, X. Q., & Liu, W. (2013). A functional polymorphism rs11614913 in microRNA-196a2 is associated with an increased risk of colorectal cancer although not with tumor stage and grade. *Biomedical reports*, *1*(5), 737-742
[41]. Milevskiy, M. J., Gujral, U., Marques, C. D. L., Stone, A., Northwood, K., Burke, L. J., & Brown, M. A. (2019). MicroRNA-196a is regulated by ER and is a prognostic biomarker in ER+ breast cancer. *British journal of cancer*, *120*(6), 621-632
[42]. Kupcinskas, J., Bruzaite, I., Juzenas, S., Gyvyte, U., Jonaitis, L., Kiudelis, G., & Kupcinskas, L. (2014). Lack of association between miR-27a, miR-146a, miR-196a-2, miR-492 and miR-608 gene polymorphisms and colorectal cancer. *Scientific reports*, *4*(1), 1-6
[43]. Zhang, H., Su, Y. L., Yu, H., & Qian, B. Y. (2012). Meta-analysis of the association between Mir-196a-2 polymorphism and cancer susceptibility. *Cancer biology & medicine*, *9*(1), 63
[44]. Vinci, S., Gelmini, S., Mancini, I., Malentacchi, F., Pazzagli, M., Beltrami, C., & Orlando, C. (2013). Genetic and epigenetic factors in regulation of microRNA in colorectal cancers. *Methods*, *59*(1), 138-146
[45]. Radanova, M., Levkova, M., Mihaylova, G., Manev, R., Maneva, M., Hadgiev, R., & Donev, I. (2022). Single Nucleotide Polymorphisms in microRNA Genes and Colorectal Cancer Risk and Prognosis. *Biomedicines*, *10*(1), 156

ISSN: 2454-9665

[46]. Tan, S. C., Lim, P. Y., Fang, J., Mokhtar, M. F. M., Hanif, E. A. M., & Jamal, R. (2020). Association between MIR499A rs3746444 polymorphism and breast cancer susceptibility: a meta-analysis. *Scientific reports*, *10*(1), 1-10.
[47]. Chu, Y. H., Tzeng, S. L., Lin, C. W., Chien, M. H., Chen, M. K., & Yang, S. F. (2012). Impacts of microRNA gene polymorphisms on the susceptibility of environmental factors leading to carcinogenesis in oral cancer. *PloS one*, *7*(6), e39777
[48]. Guo, L., Zhao, Y., Yang, S., Zhang, H., & Chen, F. (2014). Integrative analysis of miRNA-mRNA and miRNA-miRNA interactions. *BioMed research international*, *2014*
[49]. Rawla, P., Sunkara, T., & Barsouk, A. (2019). Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Przeglad gastroenterologiczny*, *14*(2), 89.