Single Nucleotide Variants Associated with Colorectal Cancer Among Iranian Patients: A Narrative Review

Mohammad Jamshidi1
Somayeh Mohammadi Pour2
Mohammad-Reza Mahmoudian-Sani3

1Department of Laboratory Sciences, School of Allied Medicine, Lorestan University of Medical Sciences, Khorramabad, Iran; 2Department of Obstetrics and Gynecology, School of Medicine Lorestan University of Medical Sciences, Khorramabad, Iran; 3Thalassemia and Hemoglobinopathy Research Center, Health Research Institute, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

Abstract: Colorectal cancer has been considered as one of the complicated multi-stage processes after adenoma-carcinoma sequence. Therefore, studies of the molecular dysregulation basis could present information on the recognition of the potent biomarkers and treatment targets for this disease. Even though outcomes of the patients with colorectal cancer have been improved largely with current annual screening plans, it is necessary to have reliable prognostic biomarkers because of the disease heterogeneity. There is a significant relationship between SNP in IL1RN2 (IL1rα), −509 C/T (TGFB1), rs11556218 T>G and rs4778889 T/C (IL16), miRNA-binding site polymorphisms in IL16, rs4464148 (SMAD7), rs6983267 (EGF), GSTT1, TACG haplotype (CTLA4), 1793G> A (MTHFR), Leu/Leu genotype of (EXO1), −137 G/C (IL18), C/T genotype (XRCC3), I3434T (XRCC7), MGMT, C3435T (MDR1), ff genotype of FokI, 677CT+TT (MTHFR), G2677T/A (MDR1) and CRC. Increased risk has been observed in VDR ApaI genotype “aa”. Finally, the protective effect has been explored in the TACA haplotype (CTLA4). According to the findings, the genetic polymorphisms in the immunity-associated genes are related to the CRC amongst the Iranian patients. Therefore, more large-scale functional investigations are necessary for confirming the results.

Keywords: single nucleotide polymorphisms, colorectal cancer, Iran, biomarker

Introduction
Colorectal cancer (CRC) is a common third leading cause of death throughout the world, which can be prevented. It is the fifth most common cancer among men and the third most common cancer among women. The disease is more common in developed countries and accounts for 65% in these countries.1 CRC has been estimated to be the fourth most common cancer in Iran.2 The prevalence of CRC in different communities varies according to different environmental factors, human behaviors, and lifestyle. Aging, inflammatory bowel disease (IBD), family history, hereditary cancer syndromes, and lifestyle-related factors (eg, diabetes, inactive lifestyle, alcohol abuse, use of the red meat, and food with low fiber) are the most important factors increasing the risk of this malignancy.3 About 80% of CRC is caused by changes that are observed as chromosomal instability, aneuploidy, and early inactivation of adenomatous polyposis coli (APC) as found in the familial adenomatous polyposis (FAP). The remaining 15% is caused by disorders that lead to microsatellite instability and defects in the function of mismatch repair (MMR) genes such as hereditary non-polyposis colorectal cancer (HNPCC).4 Moreover,
some studies found a relationship between genetic polymorphisms in individuals and an increase or decrease of their susceptibility to various cancers. Also, the genetic association studies on SNP have concentrated on the effect of single nucleotide polymorphisms on the candidate genes and cancer risk. The most important candidate genes include the genes involved in DNA repair and the immune system. SNPs can be found in coding and non-coding regions of the genes. Many SNPs do not affect the cell function; however, scientists believe that some of them can make a person susceptible to disease or affect their response to treatment. Therefore, it is difficult to find a relationship between diseases and SNPs with conventional methods because a gene may only play a minor role in the pathogenesis process. Individuals with specific SNPs or several SNPs may be more sensitive when exposed to carcinogens such as radiations. An SNP alone can increase the risk of cancer, but the presence of several polymorphic regions further increases the probability of cancer. It should be mentioned that this review has given a summary of the important genetic markers for the first time in Iranian patients with CRC (Table 1). Therefore, the mentioned genetic markers have been grouped in distinctive cellular procedures based on the respective cellular function (Figure 1).

SNPs in Genes Coding for Signaling

One of the proteins involved in the insulin signaling pathway is the insulin-like growth factor 1 (IGF1). This protein, on the one hand, stimulates cell division and, on the other hand, prevents the programmed cell death (PCD). Those with elevated IGF1 levels are 2.5 times more likely to suffer from colorectal cancer. Numerous factors such as diet, lifestyle, genetic factors, and BMI are involved in controlling the IGF1 levels. However, due to the effects of IGF1 on the cell proliferation and differentiation of the large intestine cells in vitro and the biological system and its effect on the incidence of obesity and insulin resistance, researchers investigated its role in colorectal cancer. One of the proteins binding to the IGF1 is Insulin Growth Factor Binding Protein 3 (IGFBP3), which induces the programmed cell death through two pathways of p53- and IGF1-dependent. Moreover, some studies proved that the proteins binding to the IGF1, in particular, IGFBP3, play an important role in regulating the growth of the large intestine cells and nucleotide changes in the IGFBP3 gene are involved in gastrointestinal tumors. Other studies investigated the possible role of increased vitamin D in the reduction of the risk of CRC. Vitamin D activity depends on binding to the specific intracellular vitamin D receptor (VDR), whose nucleotide sequence in the human genome is known to be a highly polymorphic gene region. The function of this receptor represents a direct relationship with the amount of vitamin D active in the circulation and the values measured in vitro. VDR protein is known to facilitate tumorigenesis in the early stages of the disease because it increases in tumor tissues with the exception of its advanced stage. It is notable that Parathyroid hormone (PTH) works as one of the main regulators of the calcium homeostasis and modifies the expression of the proteins involved in the cell cycle in the CRC cells. In this regard, some research demonstrated higher serum levels of PTH in cases with CRC in comparison to the control group. Moreover, it has been indicated that PTH gene variants had a relationship with PTH and calcium serum levels. Calcium and calcium sensing receptor (CASR) have been considered to be the structured controllers of the colonocytes and calcium via signalling across CASR, which suppressed the proliferation of the normal colonocyte. Seemingly, CASR contributed to the suppression of tumor in CRC and had an essential contribution to maintain the calcium homeostasis.

One of the studies in the field showed the fundamental involvement of epidermal growth factor (EGF) in tumor biology. EGF stimulated the cell proliferation, metastasis, invasion, angiogenesis as well as apoptosis inhibition. Also, several investigations confirmed that EGF and the respective receptor genes were over-expressed in different solid tumors, in particular, in carcinomas. Moreover, several replicative investigations proved that EGF 61A/G polymorphism could change risks for colon cancer. Another study demonstrated that EFNA1 was a glycosyl phosphatidyl inositol-linkage ligand with 205 amino acids binding to the receptor tyrosine kinase EphA2 at sites, in which the cell to cell contact occurred. The mentioned sites led to the contact-dependent bilateral signaling and played a significant role in the tumor neo-vascularization and development. Another study also showed that EFNA1 and its receptor EphA2 regulated the integrin-mediated adhesion, the cell proliferation, migration, and differentiation so that they have been regarded as the main mediators in developing and maintaining various kinds of tumors. In fact, as stated in some studies, they contributed to several oncogenesis signaling pathways, including PI3K and MAP/ERK, and may affect the initiation and development of tumors. Finally, there has been a significant relationship between EFNA1 over-expression...
Table 1 All of the Studies Markers in CRC Susceptibility Among the Iranian Patients

<table>
<thead>
<tr>
<th>SNP/gene</th>
<th>Population</th>
<th>Techniques</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL1RN VNTR, rs419598 (IL1ra)</td>
<td>91 patients 97 controls</td>
<td>PCR-RFLP</td>
<td>This study demonstrates an association between the carrier status of IL1RN* 2 and CRC [107]</td>
</tr>
<tr>
<td>-800 G/A, −509 C/T (TGFB1)</td>
<td>134 patients 138 controls</td>
<td>PCR-RFLP</td>
<td>Genetic polymorphism at −509 C/T of the TGFB1 gene may play a role in susceptibility of Iranian subjects for CRC [108]</td>
</tr>
<tr>
<td>Pro72Arg (rs1042522) (TP53)</td>
<td>100 patients 100 controls</td>
<td>ARMS-PCR</td>
<td>The alleles of the TP53 gene Pro72Arg SNP did not significantly differ in prevalence between patients and controls [109]</td>
</tr>
<tr>
<td>rs4072111, rs11556218, rs4778889 (IL16)</td>
<td>260 patients 405 controls</td>
<td>PCR-RFLP</td>
<td>Results suggest an influence of rs11556218 T &gt; G and rs4778889 T/C polymorphisms on the altered risk of CRC [110]</td>
</tr>
<tr>
<td>miRNA-binding site polymorphisms in IL16, CDKN2A (p16), RAF1, PTGER4, and ITGB4</td>
<td>249 patients 394 controls</td>
<td>PCR-RFLP</td>
<td>Significant relationship exists between the miRNA-binding site polymorphism of the IL16 gene and CRC risk [111]</td>
</tr>
<tr>
<td>rs3135500 (NOD2), rs1368439 (IL12B)</td>
<td>92 patients 105 controls</td>
<td>TaqMan assay</td>
<td>NOD2 rs3135500 and IL12B rs1368439 SNPs were not genetic risk factors for CRC [112]</td>
</tr>
<tr>
<td>rs12953717, rs4464148 in intron 3 (SMAD7)</td>
<td>234 patients 253 controls</td>
<td>PCR-RFLP</td>
<td>Significant association between CRC risk and the rs4464148 AG genotype. Although observed a strong association with rs4464148 GG genotype in affected women, did not detect the same association in CRC male patients [113]</td>
</tr>
<tr>
<td>−765G&gt;C (PTGS2)</td>
<td>110 patients 120 controls</td>
<td>PCR-RFLP</td>
<td>No significant relation was found between this polymorphism and sporadic CRC [114]</td>
</tr>
<tr>
<td>rs6983267 in the 8q24, rs4444903 (EGF)</td>
<td>115 patients 120 controls</td>
<td>PCR-RFLP</td>
<td>For the rs4444903 SNP, no significant association was found with CRC risk. 8q24 rs6983267 SNP may play a pivotal role in the development of sporadic CRC [115]</td>
</tr>
<tr>
<td>TP53 codon 72</td>
<td>250 patients 250 controls</td>
<td>PCR-RFLP</td>
<td>TP53 polymorphism and arginine/arginine genotype may be correlated with overexpression of p53 and increased risk for CRC [116]</td>
</tr>
<tr>
<td>Glutathione S-transferases P1, M1 and T1 (GSTP1, M1, T1)</td>
<td>100 patients 100 controls</td>
<td>pyrosequencing</td>
<td>GSTTI polymorphism type was significantly higher in patients. On the other hand there is no significant association between GSTM1, GSTP1 and CRC [117]</td>
</tr>
<tr>
<td>C16A genetic variation rs4359426 (CCL22) C1014T variation rs2228428 (CCR4)</td>
<td>165 patients 150 controls</td>
<td>PCR-RFLP</td>
<td>These findings collectively suggested that CCR4 C1014T and CCL22 C16A genetic variations were neither associated with the risk, nor with the progression of CRC [118]</td>
</tr>
<tr>
<td>CTLA-4 gene promoter and exon 1 (1722T/C, −1661A/G, +49A/G, 318C/T)</td>
<td>109 patients 190 controls</td>
<td>PCR-RFLP</td>
<td>A positive association between CTLA-4 TACG haplotype and CRC. A protective role for TACA haplotype is postulated [119]</td>
</tr>
<tr>
<td>rs1447295 (CASC8)</td>
<td>2416 subjects (46% patients)</td>
<td>TaqMan real-time PCR</td>
<td>No significant association between the rs1447295 polymorphism and risk of CRC [120]</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>SNP/gene</th>
<th>Population</th>
<th>Techniques</th>
<th>Results</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic variants in 8q24.21 including rs10505477 and rs6983267</td>
<td>380 patients 335 controls</td>
<td>TaqMan real time PCR</td>
<td>Results of this study suggests that the rs6983267 and rs10505477 polymorphisms alone may not be relevant to CRC risk</td>
<td>[121]</td>
</tr>
<tr>
<td>I793G&gt; A (MTHFR)</td>
<td>227 patients 239 controls</td>
<td>Pyrosequencing</td>
<td>Significant reduction in recurrence risk was seen in MTHFR G1793A heterozygotes limited to those who received folate supplements</td>
<td>[122]</td>
</tr>
<tr>
<td>L757P at exon 13 (EXO1)</td>
<td>90 patients 98 controls</td>
<td>PCR-RFLP</td>
<td>Leu/Leu genotype of EXO1 showed an inverse association with CRC.</td>
<td>[123]</td>
</tr>
<tr>
<td>−607 C/A, −137 G/C (IL1B)</td>
<td>232 patients 312 controls</td>
<td>AS-PCR</td>
<td>SNP at position −137 G/C and haplotype frequency may play a role in predisposition of Iranian patients to CRC</td>
<td>[124]</td>
</tr>
<tr>
<td>IGF-I (rs6214), IGFBP-3 (rs3110697), INSR (rs1052371), IRS2 (rs2289046)</td>
<td>167 patients 277 controls</td>
<td>PCR-RFLP</td>
<td>Not associations between polymorphic variations in IGF-I, IGFBP-3, INSR, IRS2 genes and risk of CRC</td>
<td>[125]</td>
</tr>
<tr>
<td>MGMT, DNMT1</td>
<td>208 patients 213 controls</td>
<td>PCR/pyrosequencing</td>
<td>No association between DNMT1 and CRC. However, there was a significant association between two polymorphisms in MGMT with CRC</td>
<td>[126]</td>
</tr>
<tr>
<td>C3435T MDR1</td>
<td>118 patients 137 controls</td>
<td>PCR-RFLP</td>
<td>In this study suggest that C3435T MDR1 polymorphism has an association with CRC</td>
<td>[127]</td>
</tr>
<tr>
<td>rs5277 (COX2)</td>
<td>167 patients 197 controls</td>
<td>PCR-RFLP</td>
<td>No significant difference in the distribution of COX-2 gene rs5277 polymorphism genotype and the allelic form, among CRC patients</td>
<td>[128]</td>
</tr>
<tr>
<td>rs1801725 (CaSR) rs6256 (PTH)</td>
<td>350 patients 510 controls</td>
<td>PCR-RFLP methods</td>
<td>CASR gene A986S variant is not a genetic contributor to CRC risk. PTH gene variant does not affect CRC risk</td>
<td>[129]</td>
</tr>
<tr>
<td>VDR gene ApaI and TaqI</td>
<td>160 patients 180 controls</td>
<td>PCR-RFLP</td>
<td>VDR ApaI genotype “aa” is associated with increased risk of CRC.</td>
<td>[130]</td>
</tr>
<tr>
<td>XRCC3 (T241M), XRCC3 (A17893G), XRCC7 (I3434T)</td>
<td>180 patients 160 controls</td>
<td>RFLP-PCR and ARMS-PCR</td>
<td>The incidence of CRC was observed to be significantly more in a heterozygous XRCC3 C/T genotype than in the CC genotype. XRCC7 I3434T polymorphism, CRC risk was significantly higher in I/T+T/T variant subjects compared to the I/I genotype. XRCC3 A17893G polymorphism did not correlate with CRC</td>
<td>[131]</td>
</tr>
<tr>
<td>Codon 72 of the p53 gene (Arg72Pro)</td>
<td>132 patients 163 controls</td>
<td>AS-PCR</td>
<td>Data do not support the association of the p53 codon 72 polymorphism with CRC</td>
<td>[132]</td>
</tr>
<tr>
<td>vitamin D receptor gene FokI</td>
<td>100 patients 100 controls</td>
<td>PCR-RFLP</td>
<td>FokI polymorphism may contribute to CRC susceptibility. Ff genotype of FokI polymorphism was associated with CRC risk</td>
<td>[133]</td>
</tr>
</tbody>
</table>

(Continued)
and TNM staging and the lymph node metastasis in the human gastric adenocarcinoma.\textsuperscript{27} According to the studies, SMAD7 has been considered to be a suppressive SMAD, and because of its contribution as one of the negative regulators of TGF-\(\beta\) signaling pathway, it enhances the anti-inflammatory impacts of TGF-\(\beta\) pathway.\textsuperscript{28} Hence, SMAD7 activities may notably decline TGF-\(\beta\) signaling and cause the enhanced risks of cancer.\textsuperscript{29} Concerning the genome-wide association studies (GWAS), researchers found a relationship of numerous loci and susceptibility to the CRC like diverse variants into SMAD7.\textsuperscript{30}

**Genetic Polymorphisms in the Immunity-Related Genes**

As shown in one study, the IL1 gene has been considered to be situated on chromosome 2q14, which included 3 associated genes such as IL1A, IL1B, and IL1RN that encoded IL1a, IL1b, and IL1 receptor antagonist (IL1ra). IL1ra that is one of the anti-inflammatory cytokines would competitively bind to IL1 receptors and handles the inflammatory action of IL-1.\textsuperscript{31} On the other hand, the IL1RN gene possesses an 86-bp variable number of tandem repeats (VNTR) in the second intron. Several research indicated the enhanced risk of gastric cancer\textsuperscript{32} and CRC\textsuperscript{33} by IL1RN VNTR. Another study indicated that the transforming growth factor b (TGF-b) has been a cytokine, which has been used as the tumor inhibitor in the normal intestinal epithelium via suppression of the cell proliferation and induction of apoptosis.\textsuperscript{34} Also, multiple research suggested possible contribution of abnormality in the TGF-b pathway to oncogenesis especially to the colorectal carcinoma progression.\textsuperscript{35} Another study showed that TGFB1 level has been controlled genetically and researchers observed numerous polymorphisms in the TGFB1 gene, which affected the expression of TGF-b protein.\textsuperscript{36} According to some studies, CCL22 has been considered to be a Chemokine generated basically by macrophages. Moreover, Dendritic cells and several kinds of tumor cells could be used to secrete CCL22.\textsuperscript{37,38} Also, CCR4 which has been regarded as the prominent receptor of CCL22 is a member of the G-protein coupled receptor family of the proteins. Furthermore, researchers stated this receptor expression on the surface of multiple malignant cells.\textsuperscript{38,39} According to the last reports on the immune-suppressive

<table>
<thead>
<tr>
<th>SNP/gene</th>
<th>Population</th>
<th>Techniques</th>
<th>Results</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C677T, A1298C variants of MTHFR</td>
<td>175 patients 231 controls</td>
<td>PCR-SSCP</td>
<td>MTHFR 677CT+TT variant genotype may be a risk factor for CRC</td>
<td>[134]</td>
</tr>
<tr>
<td>GT dinucleotide repeat in the PIK3CA gene</td>
<td>103 patients 150 controls</td>
<td>PCR-RFLP</td>
<td>Polymorphic GT repeat of PIK3CA gene may be a potential predictive marker of CRC</td>
<td>[135]</td>
</tr>
<tr>
<td>miR-608 (rs4919510) and miR-149 (rs2292832)</td>
<td>76 patients 70 controls</td>
<td>PCR-RFLP</td>
<td>Genotypes of rs2292832 and rs4919510 are not associated with risk of CRC.</td>
<td>[136]</td>
</tr>
<tr>
<td>SDF-1 gene at position 801 (G&gt;A)</td>
<td>109 patients 262 controls</td>
<td>PCR-RFLP</td>
<td>SDF-1 gene polymorphism at position 801 (G&gt;A) was not associated with CRC</td>
<td>[137]</td>
</tr>
<tr>
<td>C1236T, G2677T/A (MDR1)</td>
<td>60 patients 60 controls</td>
<td>PCR-RFLP ARMS-PCR</td>
<td>G2677T/A polymorphism showed association with histological grade of CRC</td>
<td>[138]</td>
</tr>
<tr>
<td>rs12904 (ephrin A1)</td>
<td>152 patients 160 controls</td>
<td>PCR-RFLP</td>
<td>No significant association between the rs12904 and sporadic CRC</td>
<td>[139]</td>
</tr>
<tr>
<td>RAD51 135G&gt;C</td>
<td>100 patients 100 controls</td>
<td>PCR-RFLP</td>
<td>RAD51 135G&gt;C probably has not a crucial role in Iranian CRC risk</td>
<td>[140]</td>
</tr>
</tbody>
</table>

Abbreviations: MGMT, O6-methylguanine-DNA methyltransferase; DNMT1, methyl transferase 1; PCR-RFLP, polymerase chain reaction–restriction fragment length polymorphism; ARMS, amplification-refractory mutation system.
contribution of CCR4 and CCL22 to numerous cancers, the CCL22 expression level has increased by gastrointestinal cells, in particular, under pathological conditions.\textsuperscript{40,41} Notably, the gene for cytotoxic T-lymphocyte antigen-4 (CTLA4) is a key gene that contributes to the immune response to different antigens. Investigations indicated the CTLA4 constitutive expression in the tumor cell lines at different intensities. Some of the also demonstrated the effects of CTLA4 gene polymorphism on the function and expression of CTLA4. Nonetheless, the most examined polymorphism of the CTLA4 gene has been reported to be an A to G substitution at position +49 in exon 1 but today researchers confirmed the effect of this SNP on the function and expression of the CTLA4 molecule.\textsuperscript{42}

Another study showed that interleukin IL18 is a proinflammatory cytokine found in an inactive precursor in the normal gut mucosa, which could be quickly converted into a biologically active molecule via interleukine-1 beta converting enzyme.\textsuperscript{43} However, studies demonstrated a relationship between multiple proinflammatory gene products like IL18 and tumorigenesis, proposing the inflammation as one of the risk factors for developing cancer.\textsuperscript{44} Hence, IL18 as a proinflammatory cytokine contributed to the gastrointestinal inflammation possibly results in progressing cancer in the gastrointestinal tract. Moreover, IL18 production and/or activity could be changed by variations in the IL18 gene promoter and thus affects the genetic susceptibility to the development of various cancers.\textsuperscript{45} It is widely accepted that SDF-1 is a CXC chemokine, which binds to CXCR7 and CXCR4 receptors and contributes to the B and T lymphocytes homing and maturation,\textsuperscript{46} angiogenesis, regulating immunity, as well as the stem cell trafficking.\textsuperscript{47} In fact, it significantly contributes to the growth, development, and metastasis of various tumors like osteosarcoma\textsuperscript{48} and breast cancer.\textsuperscript{49} SDF-1 possesses a nucleotide transition from G to A (G→A) at position 801 in the 3′-untranslated region (SDF-1-G801A) in its β transcript that is also called SDF1-3′A.\textsuperscript{50} Thus, SDF-1-3′A could have a significant modulatory contribution via enhancing SDF-1 protein generation.\textsuperscript{51} Finally, this SNP of the SDF-1 gene had a relationship to the solid tumors. Another study indicated that the gene encoding IL16 cytokine contained eight exons spanning ~17 kb of genomic DNA situated at the chromosome 15q26.3 in the human genome.\textsuperscript{52} Multiple investigations emphasized the strong relationship of inflammatory parameters like pro-inflammatory cytokines with cancers.\textsuperscript{53} Moreover, the combined inflammatory cytokines generated by the epithelial cells of the colon

**Figure 1** All the cellular processes which are studied in CRC progression among Iranian patients.
and rectum in the tumor micro-environment could importantly contribute to cancer progression.\textsuperscript{54} PTGS2 is a pro-inflammatory and inducible enzyme, which converted arachidonic acid into prostaglandins.\textsuperscript{55} In fact, induction of the PTGS2 gene expression in the epithelial cells with increased growth has been observed so that there has been a relationship with the cancer invasion and development.\textsuperscript{56} However, there are growing documents of the relationship of PTGS2 polymorphisms with the risks of CRC.\textsuperscript{57,58} Therefore, a positive relationship between SNP and the increased risks of CRC has been established about the significance of PTGS2 enzyme in the inflammatory reaction that would be one of the crucial prerequisites for developing adenoma so that this polymorphism position on the gene promoter region had a direct effect on the modulation of the gene expression and the rate of the enzyme generation.\textsuperscript{57–59}

Genetic Polymorphisms in DNA Repair-Related Genes

One of the steps to repair DNA is to remove methyl from the O-6 atom of guanine created under alkylating agents. The presence of methylated guanine nucleotides at position 6 of the oxygen atom converts the G:C base pair to A:T, which is a transition mutation.\textsuperscript{60} The enzyme O6-methyl guanine methyltransferase (MGMT) is responsible for repairing this anomaly that removes the alkyl agent from the O6 atom position of the guanine nucleotide and maintains the structure of the original genome. Non-expression of this protein increases the risk of carcinogenicity and sensitivity to the methylating agents.\textsuperscript{61} Based on the studies, MGMT methylation occurs in the promoter region occurs in normal cells that soon become cancer cells.\textsuperscript{62} Also, MGMT gene silencing with an increase in the promoter region methylation of this gene has been observed in 20–40% of patients with CRC.\textsuperscript{62–65} Exonuclease 1 (EXO1) has been considered as one of the members of the RAD2 nuclease family with a significant contribution to the DNA replication, recombination, and mismatch repair.\textsuperscript{66} Also, the probable functional polymorphisms in EXO1 could involve in changing the CRC risks via effects on the repair activities of EXO1.\textsuperscript{67} Thus, this hypothesis would be reasonable that there is possibly a relationship between SNPs in EXO1 and CRC risks.

It has been also shown that RAD51 would play a significant role in the double-strand breaks (DSB) repair of DNA. In fact, the SNPs in this gene may affect the DNA repair potential and consequently the susceptibility to different tumors like CRC. In addition, RAD51 crucially contributes to the DSB repair across the homologous recombination (HR). It is notable that a major pathway to the DNA repair is the excision repair (ER) and double-strand break repair (DSBR). In fact, DSBR contained 2 mechanisms of homologous recombination (HR) and nonhomologous end joining (NHEJ). Therefore, different proteins and enzymes like XRCC3 and DNA-dependent protein kinase (DNA-PK) involve in each mechanism for repairing the injured DNA. Actually, the protein encoded by XRCC3 gene has been XRCC3 and DNA-PKcs has been the product of XRCC7 gene. Notably, the two proteins contributed to the DSBR mechanism; that is, DNA-PK in NHEJ and XRCC3 in HR.\textsuperscript{68} Recent researchers investigated the relationship between CRC and X-ray repair cross-complementing protein 3 (XRCC3) as the gene involved in the homologous recombination pathway.\textsuperscript{69–71} XRCC3 is an essential protein for chromosomal stability and cellular resistance to radiation and some chemical agents; however, in spite of its importance in repairing DNA. DSBs through homologous recombination pathway, there is not enough information of its biochemical properties and specific function.\textsuperscript{72} As a basic origin of dietary methyl groups, folate contributes importantly to the DNA methylation, repair, and synthesis. In addition, 5.10-methylene-tetrahydrofolate reductase (MTHFR) has been considered to be a major enzyme in the folate metabolism that performs irreversible conversion of 5.10-methylene-tetrahydrofolate into 5-methyl-tetrahydrofolate that would direct the folic acid pool towards re-methylation of homocysteine to methionine.\textsuperscript{73} It is notable that folate deficiency can result in the uracil mis incorporation and consequent DNA instability,\textsuperscript{74} retarded DNA repair capacities for oxidative or alkylating damage,\textsuperscript{75} and global and proto-oncogenic DNA hypomethylation.\textsuperscript{76} Thus, each effect contributes to carcinogenesis and the increased intake of folate has a relationship with lower risks of a number of cancers like CRC.\textsuperscript{77} Nonetheless, the relationship between folate and CRC has been incompatible with regard to the last studies. Moreover, there is a relationship between sufficient folate intake and the highly declined risk of CRC,\textsuperscript{78} however, there has been not a relationship with the endometrial cancer risks.\textsuperscript{79}

SNPs in the Cell Cycle Genes

Some studies recognized approximately 14 SNPs in wild type TP53(p53) gene that can alter the p53 protein
function. Pro72Arg (rs1042522) is a common SNP of the TP53 gene situated at the proline-rich domain of p53, which is crucial for normal p53 functions. Moreover, arginine (Arg) variant could induce apoptosis more rapidly and efficiently than proline (Pro) whereas Pro variant had a more acceptable function to induce the cycle arrest. The cancer risk could be enhanced by Pro72Arg SNP in the TP53 gene. One of the studies in the field showed the relationship of IL-16, CDKN2A (p16), RAF1, PTGER4, and ITGB4 with various cancers. CDKN2A has been considered to be a popular gene due to the respective impacts on pancreatic cancer. Moreover, ITGB4 has been categorized into the integrin protein group. The above proteins should provide the grounds for the cell-cell and cell-extracellular matrix adhesion. Because of such properties, the above proteins could contribute to cancer development. PTGER4 works as one of the negative feedback regulators of cell proliferation or rapid growth and thus the respective changes may cause the tumor progression. Therefore, variations in expressing the cancer-associated genes, CDKN2A (p16), RAF1, PTGER4, and ITGB4 amongst a population could apply negative or positive impacts on the individuals’ susceptibility to cancers.

Non-Coding RNAs and Epigenetic Modification in CRC

According to a study in the field, the genetic polymorphism in the miRNA-binding region of the mRNA’s 3’UTR could lessen the miRNA-mRNA interactions, change expressing the target gene, and influence the individuals’ risks of disease. Numerous investigations demonstrated a potent contribution of rs1447295 polymorphism to susceptibility to cancers. This variant situated at the cancer susceptibility candidate 8 (CASC8) has been considered to be a long non-coding RNA (lncRNA) gene that would not code protein. CASC8 has been located adjacent to the Myc gene in the 8q24.1 region that is a popular gene desert consisting of several enhancer elements in the proximal of the MYC gene, associated CRC. Finally, the above enhancers modulate the MYC gene transcription via interacting with the CASC8 promoter.

Polymorphisms of Anti-Oxidant Enzymes in CRC

Glutathione S-transferases (GSTs) that has been considered as one of the superfamilies’ of the dimeric Phase II metabolic enzymes contribute importantly to the cellular defense mechanism. Research showed a wide expression of sub-class GSTP1 in the normal human epithelial tissues and high overexpression in colon cancer. Moreover, GSTM1 and GSTT1 deficiency is related to higher risks of specific cancers. The epidemiologic investigations studying the relationship between variants in CYP and GST genes and colorectal neoplasia presented inconsistent outputs and thus there has been no agreement on their etiologic significance. However, a way to examine the protective contribution of GSTs would be to study the polymorphism in GSTM1, GSTT1, GSTP1, and CYP2E1 genes on the susceptibility to the CRC.

Microsatellite Instability in CRC

In 1993, it was found that the dysfunction of genes involved in the mismatch repair (MMR) pathway leads to microsatellite instability (MSI) in tumors with this defect. The cause of MSI is the presence of a mutation in the germline of the MMR system. The microsatellite sequence is scattered throughout the genome and consists of tandemly repeated sequences of 1–6 base pairs, which show very high polymorphism. In some tumors, half or more of the microsatellite become instable; a condition known as high-level microsatellite instability or MSI-High. Depending on the degree of instability, MSI is classified into three classes of MSI-High (MSI-H), MSI-Low (MSI-L), and MS-Stable (MSS). Microsatellite instability measurement is an excellent and fairly easy way to detect the deficiency of proteins of the MMR system. Among the molecular markers considered for the diagnosis and characterization of colorectal cancer, MSI has the advantage of providing promising information about the recurrence of this cancer. Colorectal cancer...
patients with MSI-High have multiple deletion/insertion mutations in at least two of the five DNA loci.\textsuperscript{106}

Conclusions
Finding new prognostic factors, new biomarkers or pathologically modified characteristics is a never ending story and would be the subject of future research. Also, a personal approach to medicine (personalized medicine) has changed the field of oncology during the last decade. Tumor, Node, Metastasis (TNM) clinical and pathologic staging system is still the most important prognostic factor; however, in certain cancers and specific stages, it does not provide enough prognostic information and predictive information due to the heterogeneity of CRC. Therefore, new molecular objectives lead to the study of the biomarkers as the prognostic and predictive agents. Understanding the molecular mechanisms involved in the CRC development and metastatic stages will help us identify people at the highest risk of recurrence and find new tumor targets to prevent disease progression. Thus, one of the new biomarkers in CRC may be SNPs that affect the disease incidence and have the potential to be a prognostic or predictive agent for daily clinical actions and decisions. However, it is important to identify the best and most relevant SNPs. According to Table 1, the presence of SNP in $-800$ G/A (TGFBI), Pro72Arg rs1042522 (TP53), rs4072111 (IL16), miRNA-binding site, CDKN2A (p16), RAFl, PTGER4, ITGB4, rs3135500 (NOD2), rs1368439 (IL12B), rs12953717 (SMAD7), $-765$G$>C$ (PTGS2), rs4444903 (EGF), GSTM1, GSP1, rs4359426 (CCL22), rs2228428 (CCR4), rs1447295 (CASC8), rs6983267 and rs10505477 (enetic variants), $-607$ G/A (RAD51), C17893G (XRCC3) genes has no significant relationship with any DNA loci.

Table 1
<table>
<thead>
<tr>
<th>SNP Position</th>
<th>Gene</th>
<th>Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-800$ G/A</td>
<td>TGFBI</td>
<td>Pro72Arg</td>
</tr>
<tr>
<td>rs1042522</td>
<td>TP53</td>
<td>rs72Arg</td>
</tr>
<tr>
<td>rs4072111</td>
<td>IL16</td>
<td>miRNA-binding site</td>
</tr>
<tr>
<td>CDKN2A(p16)</td>
<td>RAFl</td>
<td>PTGER4</td>
</tr>
<tr>
<td>ITGB4</td>
<td>rs3135500</td>
<td>NOD2</td>
</tr>
<tr>
<td>rs1368439</td>
<td>IL12B</td>
<td>rs12953717</td>
</tr>
<tr>
<td>rs4444903</td>
<td>PTGS2</td>
<td>rs4444903</td>
</tr>
<tr>
<td>rs4359426</td>
<td>CCL22</td>
<td>rs4359426</td>
</tr>
<tr>
<td>rs2228428</td>
<td>CCR4</td>
<td>rs2228428</td>
</tr>
<tr>
<td>rs1447295</td>
<td>CASC8</td>
<td>rs1447295</td>
</tr>
<tr>
<td>rs6983267</td>
<td>rs10505477</td>
<td>enetic variants</td>
</tr>
<tr>
<td>rs10505477</td>
<td>RAD51</td>
<td>C17893G</td>
</tr>
</tbody>
</table>

Future Perspectives
Molecular pathological epidemiology (MPE) is an integrative discipline that combines epidemiology with molecular pathology. This field is structured around the core principle that diseases can be viewed as the product of profiles of exposomes, epigenomes, transcriptomes, proteomes, metabolomes, microbiomes, and interactomes and how they affect and are affected by macroenvironment and tissue microenvironment. This core principle, which is known as the unique disease principle, distinguish MPE from genome-wide association studies (GWAS).\textsuperscript{141} MPE helps researchers to gain valuable insights into the heterogeneity of diseases and produce epidemiologic data to further the study of molecular pathogenic mechanisms. With the wealth of biomedical data available to today’s researchers, they can use MPE to refine these data to gain a better understanding of the etiology and pathogenesis of diseases.\textsuperscript{142} MPE studies of CRC premalignant lesions can greatly contribute to personalized prevention, screening, and treatment of CRC cancer by offering valuable insights into the etiological factors of neoplastic initiation and progression and therefore the causes and risk factors of this cancer.

Ethical Approval
This article does not contain any studies with human participants or animals performed by any of the authors.

Disclosure
The authors report no conflicts of interest in this work.

References


