MINI-REVIEW

Genes and SNPs Associated with Non-hereditary and Hereditary Colorectal Cancer

Mohammadreza Nassiri¹,²*, Mohammad Mahdi Kooshyar³, Zahra Roudbar¹, Morteza Mahdavi¹, Mohammad Doosti¹

Abstract

Background: Colorectal cancer is the third most common cancer in both men and women in the world and the second leading cause of cancer-related deaths. The incidence of colorectal cancer has increased in Iran in the past three decades and is now considered as a serious problem for our society. This cancer has two types hereditary and non-hereditary, 80% of cases being the latter. Considering that the relationship between SNPs with diseases is a concern, many researchers believed that they offer valuable markers for identifying genes responsible for susceptibility to common diseases. In some cases, they are direct causes of human disease. One SNP can increase risk of cancer, but when considering the rate of overlap and frequency of DNA repair pathways, it might be expected that SNP alone cannot affect the final result of cancer, although several SNPs together can exert a significant influence. Therefore identification of these SNPs is very important. The most important loci which include mutations are: MLH1, MSH2, PMS2, APC, MUTYH, SMAD7, STK11, XRCC3, DNMT1, MTHFR, Exo1, XRCC1, and VDR. Presence of SNPs in these genes decreases or increases risk of colorectal cancer.

Materials and Methods: In this article we reviewed the Genes and SNPs associated with non-hereditary and hereditary of colorectal cancer that recently were reported from candidate gene y, meta-analysis and GWAS studies.

Results: As with other cancers, colorectal cancer is associated with SNPs in gene loci. Generally, by exploring SNPs, it is feasible to predict the risk of developing colorectal cancer and thus establishing proper preventive measures.

Conclusions: SNPs of genes associated with colorectal cancer can be used as a marker SNP panel as a potential tool for improving cancer diagnosis and treatment planning.

Keywords: Colorectal cancer - hereditary genes - non-hereditary genes - SNPs

Introduction

Colorectal cancer is the third most common cancer in both men and women and the second most common cause of cancer-related deaths in the world (Brenner and Rosenberg, 2010). There are nearly one million new cases of diagnosed CRC and half a million deaths worldwide each year and its incidence has shown a growing process in the last century (Azadeh et al., 2008). CRC develops sporadically most frequently at an increasing rate in young patients (<50 years), although the highest prevalence still occurs in those aged 60-70 years (Berg and Soreide, 2011). Data which is available from CRC incidence showed a remarkable increase in past three decades in Iran. Thus, the importance of CRC as a public health problem is increasing in Iran (Moghimi-Dehkordi et al., 2008).

This cancer has two types, hereditary (familial) and non-hereditary (sporadic) forms, in which 80% of cases are as non-hereditary, while the remaining 20% have an inheritance pattern (Alvarado et al., 2006; Gopalan et al., 2010). Single nucleotide variations in genomic sequences are called SNPs. The SNP is just a single base change in a DNA sequence, with a usual alternative of two possible nucleotides at a given position. Approximately, there are 10 million SNPs in the human genome and current estimates suggest that there is an average one common SNP in every 300 base pairs throughout the human genome. At least 5 million SNPs have been reported in SNP databases.

SNPs may fall within coding regions, non-coding regions or in intergenic regions (Vali et al., 2008). SNPs in gene coding regions can lead to changes in the biological properties of the encoded protein. In contrast, SNPs in non-coding gene regulatory regions may affect gene expression levels in an allele-specific manner and these functional SNPs represent an important but relatively unexplored class of genetic variation (Wang et al., 2004). Researchers believe that SNP map will help in finding effective genes in complex diseases such as cancer. Identification of SNPs’ sequences which regulate gene expression is a key to understanding human genetic differences and diseases (Lam et al., 2001; Salajegheh et al., 2011). Recently, a
lot of research on the gene expression and polymorphism causing various cancers has been done that indicates the importance of identifying SNPs in cancer (Smith et al., 2011). Given that no review article has been published about genes and SNPs associated with non-hereditary and hereditary of CRC, our studies on some genes associated with CRC and works that have been done by other researchers prompted us to here in review genes and SNPs associated with non-hereditary and hereditary of CRC.

Genes Associated with Non-Hereditary Colorectal Cancer

**XRCC3** gene

X-ray repair cross-complementing protein 3 (XRCC3) gene is located on chromosome 14q23.3 and encodes a member of the RecA/Rad51 related protein family that is involved in the repair of DNA double strand breaks in the process of homologous recombination in order to maintain chromosome stability (http://www.ncbi.nlm.nih.gov/gene). Chromosomal instability and impaired cell viability have been correlated with mutations in XRCC3. The correlation with chromosomal instability indicates that this may be a good candidate for a tumor suppressor gene. The C18067T substitution in XRCC3 in exon 7 is a non-conservative change, but does not reside in the functional ATP-binding domains (Ladiges et al., 2003).

Jin et al. (2005) studied the relationship between XRCC3 polymorphism in position C18067T with the risk of CRC in a Chinese population and reported that the C18067T polymorphism in the homologous recombination repair gene XRCC3 may alter DNA repair capacity and subsequent susceptibility to carcinogens. The C18067T nucleotide substitution in exon 7 is the most frequent polymorphism in XRCC3, which may affect the coding enzyme’s function and/or its interaction with other proteins involved in DNA repair. In their study, they used PCR-RFLP method for genotyping and reported that the genotypes are in Hardy-Weinberg equilibrium and are associated with the risk of CRC (Jin et al., 2005). A study was conducted to investigate the effect of genetic polymorphisms of DNA repair genes XRCC1 and XRCC3 and risk of CRC in Chinese population. This study was conducted with 485 cases and 970 controls and genotyping was performed using PCR-CTPP (Polymerase chain reaction with the confronting two pair primer) method and showed that allele C of XRCC3 gene is moderately associated with CRC risk and heavy cancer risk was found in rectum cancer (Zhao et al., 2012).

**DNMT1** gene

DNA Methyl-1 enzyme is called DNA Methyle Transfase 1 gene (DNMT1 gene) in human. This gene is located on 19p13.2 chromosome with 61734 bp length and encodes a protein with 1616 amino acids. DNMT1 has a role in the establishment and regulation of tissue specific patterns of methylated cytosine residues. Aberrant methylation patterns are associated with certain human tumors and developmental abnormalities. Two transcript variants encoding different isoforms have been found for this gene (http://ghr.nlm.nih.gov/gene/DNMT1).

Kanai et al. (2003) investigated the mutation of DNMT1 1 gene in human CRC by PCR-SSCP analysis using 46 oligonucleotide primer sets for all 40 coding exons and the 50-flanking region (450 bp) of the DNMT1 gene in 29 CRC patients and showed that mutations in coding exons of the DNMT1 gene were detected in two (7%) of the 29 CRC cases. This mutation consisted of a one base deletion at position 2389 in exon 23 resulting in a premature stop codon at positions 2564-2566 and a point mutation (A to G) at position 4441 in exon 35 (Kanai et al., 2003). They suggested that mutational inactivation of the DNMT1 gene that potentially causes a genome-wide alteration of DNA methylation status may be a rare event during human carcinogenesis (Kanai et al., 2003). A study was performed to determine the relationship between SNPs (rs61750053’’ rs62621087’’ rs16999358’’ rs61750052’’ rs16666593 and rs2228613) in DNMT1 gene and CRC in patients who were referred to Taleghani hospital in Tehran. They observed that all investigated SNPs showed significant relationships with CRC, indicating that in all cases the chance of getting CRC in people with dominant genotypes was much higher than those with other genotypes. By identifying SNP sequence, it is feasible to predict the risk of catching CRC and thus establishing proper preventive measures (Alipour-Heidari et al., 2011).

**MTHFR** gene

Methylenetetrahydrofolate reductase (MTHFR) is located on 1p36 and coding 5,10 methylenetetrahydrofolate reductase enzyme. This gene has important role in the relationship between construction cycle and DNA methylation. MTHFR is the key enzyme in the metabolism of folate, carrying out the irreversible conversion of 5,10-methylenetetrahydrofolate to 5methyltetrahydrofolate that is the first substance required as a substrate for DNA synthesis while the second substance is involved in DNA methylation. MTHFR gene is a polymorphic gene; there are some variants of SNPs in this gene. A common SNP in MTHFR (C677T) is associated with decreased enzyme function in vitro and has been implicated in the risk of developing CRC (http://ghr.nlm.nih.gov/gene/MTHFR).

Haghighi et al. (2008) conducted a study to investigate the association between the G1793A MTHFR polymorphism and sporadic CRC in Iran. In this study, SNPs were identified using Pyrosequencing method and showed that the G1793A genotype of MTHFR gene is inversely associated with the risk of sporadic CRC in different populations. This means that people with GA genotype have a protective ability against sporadic CRC (Haghighi et al., 2008). Also, a case-control study was performed to evaluate the potential role of the methylenetetrahydrofolate reductase (MTHFR) C677T gene polymorphism in CRC and reported that the MTHFR 677TT genotype was significantly correlated with increased risk of CRC. This study suggests that the MTHFR C677T polymorphism indicates susceptibility to CRC and is correlated with CRC pathogenesis, suggesting that the homozygous variant MTHFR C677T polymorphism is a candidate risk factor for CRC (Yin et al., 2012).
Base excision repair (BER) is the predominant DNA damage repair pathway for the processing of small base lesions derived from oxidation and alteration damage. X-ray repair cross complementing group 1 (XRCC1) is one of the most important proteins in BER and is closely associated with BER pathway coordination by interacting with most components of the BER short patch pathway. The gene is mapped to chromosome 19q13.2-13.3 and consists of 17 exons. It encodes a 2.2 kb transcript which corresponds to a putative protein of 633 amino acids. The XRCC1 gene exhibits more than 300 SNPs, among which approximately 35 variants are located in the exons or the promoter regions of the gene. A growing number of reports on XRCC1 SNPs has been shown to be associated with measurable reduced DNA repair capacity and increased the risk of several types of cancers including colon cancer (Jelonek et al., 2010).

Rowyda et al. (2011) investigated the possible association of three single nucleotide polymorphisms C194T, G399A, and G280A, in XRCC1 gene with the risk of developing CRC in Saudi patients. To determine the polymorphic alleles, XRCC1 C194T, G280A and G399A PCR products were digested with MspI, Rsal and MspI enzymes respectively. The authors reported that the G280A polymorphism of XRCC1 DNA repair gene may contribute to genetic susceptibility to CRC and G399A may have a protective role in decreasing CRC risk and suggested that genetic polymorphisms in the XRCC1 gene (G280A and G399A) may play an important role in the development of CRC in the Saudi population (Rowyda et al., 2011).

**Exo1 gene**

Exonuclease 1 (Exo1) gene is a candidate gene for colorectal tumor susceptibility. Exo1 gene is an important nuclease involved in the MMR system that contributes to the maintenance of genomic stability, modulation of DNA recombination and mediation of cell cycle arrest. Potential polymorphisms in Exo1 may alter cancer risks by influencing the repair activity of Exo1. The Exo1 gene is a member of the MMR system and also belongs to the RAD2 nuclease family. It is located at chromosome 1q42-q43 and contains one untranslated exon followed by 13 coding exons and encodes an 846-amino acid protein (Yamamoto et al., 2005). On the other hand, Exo1 is the only enzyme known to exist in the MMR system that helps to maintain genomic stability in human and which has both endonuclease and exonuclease activities (Tsai et al., 2009).

Haghhighi et al. (2011) investigated the association between C>T polymorphism in Exo1 gene and risk of sporadic CRC in an Iranian population. This polymorphism is located on exon 13 and genotype analysis results showed that people with T/T genotype had less risk of CRC and gained a protective resistance against sporadic CRC, while C/T genotype did not reveal a significant association with CRC at the same status (Haghhighi et al., 2011). In addition, despite other investigations, they could define a significant association between T allele and colorectal cancer (Haghhighi et al., 2011).

**XRCC1 gene**

Base excision repair (BER) is the predominant DNA damage repair pathway for the processing of small base lesions derived from oxidation and alteration damage. X-ray repair cross complementing group 1 (XRCC1) is one of the most important proteins in BER and is closely associated with BER pathway coordination by interacting with most components of the BER short patch pathway. The gene is mapped to chromosome 19q13.2-13.3 and consists of 17 exons. It encodes a 2.2 kb transcript which corresponds to a putative protein of 633 amino acids. The XRCC1 gene exhibits more than 300 SNPs, among which approximately 35 variants are located in the exons or the promoter regions of the gene. A growing number of reports on XRCC1 SNPs has been shown to be associated with measurable reduced DNA repair capacity and increased the risk of several types of cancers including colon cancer (Jelonek et al., 2010).

Rowyda et al. (2011) investigated the possible association of three single nucleotide polymorphisms C194T, G399A, and G280A, in XRCC1 gene with the risk of developing CRC in Saudi patients. To determine the polymorphic alleles, XRCC1 C194T, G280A and G399A PCR products were digested with MspI, Rsal and MspI enzymes respectively. The authors reported that the G280A polymorphism of XRCC1 DNA repair gene may contribute to genetic susceptibility to CRC and G399A may have a protective role in decreasing CRC risk and suggested that genetic polymorphisms in the XRCC1 gene (G280A and G399A) may play an important role in the development of CRC in the Saudi population (Rowyda et al., 2011).

**VDR gene**

Vitamin D is a steroid hormone that its biological function is through the metabolite 1α, 25 dihydroxy vitamin D3 active. VDR belongs to the nuclear hormone receptor superfamily and acts as a ligand inducible transcription factor. The vitamin D receptor (VDR gene) is a crucial mediator of the cellular effects of vitamin D and additionally interacts with other cell signaling pathways that influence cancer development (Kato, 2000). The gene encoding VDR is located on chromosome 12q13, contains 11 exons, spans approximately 75 kilobases of genomic DNA, and has an extensive promoter region capable of generating multiple tissue specific transcripts (http://www.ncbi.nlm.nih.gov/gene). Several single nucleotide polymorphisms have been identified that may influence cancer risk. The most frequently studied single nucleotide polymorphisms are the restriction fragment length polymorphisms ApaI, BsmI, FokI and BsmI, as defined by the endonucleases. The FokI restriction fragment length polymorphism is located in the coding region of the vitamin D receptor (VDR gene). The TaqI polymorphism is a T/C nucleotide substitution (ATT to ATC) leading to a synonymous change at codon 352 in exon IX. BsmI and ApaI restriction site polymorphisms occur in the intron separating exons VIII and IX. VDR gene has also been suggested as one of the candidate genes for genetic control of CRC and vitamin D receptor (VDR) gene variants have been variably associated with the risk of CRC in epidemiological studies (Uitterlinden et al., 2004; Ochs-Balcom et al., 2008).

Dilmec et al. (2009) studied the genotype and allele frequencies and association of the VDR gene BsmI G>A (rs1544410) polymorphism with CRC in Şanlıurfa population. They are tested on 56 patients with CRC from Şanlıurfa province, and controls were randomly selected from 169 healthy individuals from the same area. Their DNA was isolated from whole blood and scanned using PCR-RFLP to determine the frequencies of BsmI G>A polymorphism of VDR gene. Results showed that VDR gene BsmI G>A allele increased the risk for CRC. After studying the association of allele 61968C>T (rs731236) polymorphism in VDR gene and CRC (Dilmec et al 2009).

**Genes Associated with Hereditary Colorectal Cancer**

**MLH1 gene**

MutL homolog 1 (MLH1) is the key component of the MMR system which participates in the recognition of nucleotide mismatches occurring during DNA replication and in the recruitment of additional mismatch repair proteins to the site to correct the replication error. This gene has a length of 57375 kb with 19 exons that encodes a protein with 756 amino acids and is located on the short (p) arm of chromosome 3 at position 21.3 (Hampel et al., 2005). Three single nucleotide polymorphisms in this gene are of particular interest because of their prevalence and potential to affect mismatch repair functions. These SNPs are located in the MLH1 gene (intronic IVS14-19A>G SNP, -93G>A promoter SNP and 1219V coding SNP) in the intronic IVS14-19A>G SNP located 19 nucleotides
are associated with the risk of CRC. SNP (118T>C) polymorphisms in DNA MMR genes MSH2 and MSH6 were identified as a locus frequently mutated in hereditary mismatches, thereby initiating DNA repair. MSH2 was a beta (MSH2-MSH3 heterodimer) which bind to DNA MutS alpha (MSH2-MSH6 heterodimer) and MutS homolog 2 (MSH2) is a gene commonly associated with hereditary colorectal cancer. This gene is located on chromosome 2p21. Components of the post replicative DNA repair. The oxidized 8-hydroxyguanosine gene encodes a DNA glycosylase involved in oxidative damage repair. The oxidized 8-hydroxyguanosine was detected using the fluorogenic 5’ nuclease assay. They observed strong associations between MSH2 118T>C polymorphism and family history of CRC. This association was especially evident among female CRC patients. It is possible that this SNP plays a role in the MSH2 response to sex hormones (Mrkonjic et al., 2007).

**PMS2 gene**

This gene is one of the PMS2 gene family members which are found in clusters on chromosome 7. Postmeiotic segregation increased 2 (PMS2 gene) is located from 6,012,869 to 6,048,736 base pair on chromosome 7 that encodes a protein with 862 amino acids. The product of this gene is involved in DNA mismatch repair and forms a heterodimer with MLH1. This complex interacts with MSH2 bound to mismatched bases. Defects in this gene are associated with hereditary nonpolyposis colorectal cancer (http://www.ncbi.nlm.nih.gov/gene/5395). Hazra et al. (2008) performed a large scale evaluation of genetic variants in candidate genes for CRC risk in the nurses. They reported that PMS2 is a key gene in the MMR pathway. The PMS2-24G>C SNP (rs6463524) on exon7 was associated with an increased risk of CRC (Hazra et al., 2008).

**APC gene**

This gene is a major gene that is involved in causing hereditary colorectal cancer. APC is classified as a tumor suppressor gene. Tumor suppressor genes prevent the uncontrolled growth of cells that may result in cancerous tumors. The protein made by the adenomatous polyposis coli (APC gene) plays a critical role in several cellular processes that determine whether a cell may develop into a tumor. The APC protein helps to control how often a cell divides, how it attaches to other cells within a tissue, or whether a cell moves within or away from a tissue. This protein also helps to ensure that the chromosome number of cells produced through cell division is correct (Narayan et al., 2007). The human APC gene is located on the long (q) arm of chromosome 5 between positions 21 and 22, from 112,118,468 to 112,209,532 base pair, contains 15 exons, and has a corresponding mRNA of approximately 10 kb (Ionescu et al., 2005). Chen et al. (2006) conducted a study to evaluate association between SNPs of the APC gene and CRC risk in a Taiwanese population. To compare the genotype distribution of variant sites, the full-length APC genes of 74 healthy individuals and 80 CRC patients were sequenced with ABI3100 Genetic analyzer and 12 SNPs were identified. Of these variants, 8 SNPs were located within exon 15, 2 SNPs within exon11, 1 SNP within exon 9, and 1 SNP within exon 13. reported that only three of these SNPs (T139637C, G 152553A, T153342G) were associated with the increased risk of CRC (Chen et al., 2006).

**MUTYH gene**

The mutY homolog (MUTYH gene) is located on the short (p) arm of chromosome 1 between positions 34.3 and 32.1, from 45,464,007 to 45,475,152 base pairs. This gene encodes a DNA glycosylase involved in oxidative DNA damage repair. The oxidized 8-hydroxyguanosine
(8-oxo-G) has the ability to mispair with adenine, and results in G:C to T:A and A:T to C:G transversions after replication. MUTYH enzyme is a member of the mammalian BER glycosylases that excise adenines misincorporated opposite to 8-oxo-G. The protein is localized to the nucleus and mitochondria. Mutations in this gene result in a heritable predisposition to colon and stomach cancers (Tsuzuki et al., 2007). Tao et al. (2008) studied the association between genetic polymorphisms of MUTYH gene and increased CRC risk. They investigated four MUTYH SNPs, IVS1+11C>T, IVS6+35G>A, IVS10-2A>G, and 972G>C, for association with increased CRC risk in a population based series of 685 CRC patients and 778 control subjects from Kyushu, Japan. They observed a statistically significant association between IVS1+1IT and increased CRC risk. This polymorphism occurs in the region between exon 1 and intron 1 of MUTYH gene. Also, to investigate the association of MUTYH G 13638 C and APEX1 T6865G with CRC and smoking in a Japanese population (Tao et al., 2008), Kasahara et al. (2008) conducted a study. In this study the genetic polymorphisms of DNA repair enzymes MUTYH G 13638 C and APEX1 T6865G were determined using PCR-RFLP and results showed that the MUTYH G 13638 C is strongly associated with CRC susceptibility. For this reason, if this polymorphism is in people, there is the risk of CRC (Kasahara et al., 2008).

**SMAD7 gene**

SMAD7 is a protein in human that is encoded by the SMAD7 gene. It belongs to the SMAD family of proteins, which belongs to the TGFβ superfamily of ligands. SMAD7 is an inhibitory SMAD and a negative regulator of the TGF-β signaling pathway that promotes the anti-inflammatory effects of TGF-β signaling via binding to TAB2 and TAB3 and inhibiting TAK1 (Hong et al., 2007). The mothers against decapentaplegic homolog 7 (SMAD7 gene) is located on the long (q) arm of chromosome 18 at position 21.1, from 46,446,222 to 46,477,080 base pair. Although SMAD7 has been shown to induce hepatic metastasis in CRC, its role in cancer development, particularly CRC is not fully described. Variations in this gene are the causes of susceptibility to CRC. Two recent genome-wide association studies (GWAS) identified three common variants in SMAD7 (rs4464148, rs4939827 and rs12953717) that confer modest susceptibility to CRC (Tenesa et al., 2008).

Martha et al. (2010) evaluated SNPs in the SMAD7 gene, including rs4939827, rs12953717, and rs4464148, previously identified from GWAS in a large population-based case-control study of the CRC. They observed that rs12953717 was associated with a statistically significant increased risk of CRC whereas the rs4939827 SNP was inversely associated with CRC (Martha et al., 2010).

**STK11 gene**

Serine/threonine protein kinase 11 (STK11) also known as liver kinase B1 (LKB1) or renal carcinoma antigen NY-REN-19 is a protein kinase that is encoded by STK11 gene and located on the short (p) arm of chromosome 19 at position 13.3, from 1,205,797 to 1,228,433 base pair. STK11 consists of nine exons. STK11, a tumor suppressor gene which means it, helps to keep cells from growing and dividing too fast or in an uncontrolled way. In addition to its role in regulating cell division, this enzyme helps certain types of cells to orient themselves correctly within tissues and assists in determining the amount of energy use of a cell (Grahame, 2003).

Martha et al. (2010) investigated genetic variations in nine important genes which belong to metabolic signaling pathways and CRC risk including: mTOR, PTEN, STK11, RPKAA1, PRKAG2, TSC1, TSC2, PI3K and Akt1. They reported that the rs741765 polymorphism in STK11 gene is obtained by the replacement of G/A, and is associated with CRC (Martha et al., 2010).

**Conclusion**

SNPs in the MLH1, MSH2, PMS2, APC, MUTYH, SMAD7, STK11, XRCC3, DNMT1, MTHFR, Exo1, XRCC1 and VDR genes have been associated with decreased or increased risk of CRC. Thus identification of SNPs is of major importance and of course SNPs of gene associated with hereditary CRC are suitable for screening CRC in progeny, and SNPs of gene associated with nonhereditary CRC are used as an SNP marker which is a potential tool for improving cancer diagnosis and treatment planning. Recently, the cost and the time required for sequencing are reduced, and this provides a good opportunity for screening all genes associated with CRC risk, and because of the cost reduction in next generation whole genome sequencing procedures may allow discovery of SNPs associated with CRC. Generally, by exploring people’s SNPs, it is feasible to predict the risk of catching CRC and thus establishing proper preventive measures.

**References**


