Polymorphism (rs16917496) at the miR-502 Binding Site of the Lysine Methyltransferase 5A (SET8) and Its Correlation with Colorectal Cancer in Iranians

Abstract

Background: One of the gene expression regulatory mechanisms is mediated by small noncoding RNAs called microRNA (miRNA). They interact with a recognition sequence located mostly in 3′-untranslated regions (3′-UTRs) of mRNAs. Polymorphisms in miRNAs recognition sequences could affect gene expression which in turn may alter disease susceptibility. SET8, a member of the SET domain-containing methyltransferase, acts in a variety of biological processes such as genomic stability. Here, we report correlation of rs16917496 polymorphism, located in the recognition sequence of miR-502 within 3′-UTR of SET8, with colorectal cancer (CRC) in Iranians. Materials and Methods: One hundred and seventy CRC patients and 170 noncancer counterparts were recruited in this case–control study. Genotyping of rs16917496 was performed using polymerase chain reaction-restriction fragment length polymorphism method. Results: There was no significant association of rs16917496 with CRC in population under study (P value for genotype and allele distribution were >0.05). However, stratification analysis based on smoking status revealed that TT+TC genotypes of SET8 rs16917496 are strongly associated with increased risk of CRC (odds ratio: 5.8, 95% confidence interval: 1.37–24.34, P = 0.005) in smoker subgroup. Conclusion: Correlation of rs16917496 T allele with CRC in smokers is emphasizing the importance of individuals' genotype in the recruitment of adverse health hazards of smoking more profoundly for certain people compared to others.

Keywords: Colorectal cancer, microRNA, SET8 gene, single-nucleotide polymorphism

Introduction

Colorectal cancer (CRC) is the second most common cancer in women and the third prevalent one in men in developed countries.[1-3] According to the National Institute of Cancer Statistics as of 2013, 45 out of 100,000 people diagnosed with CRC in the world.[4,5] Although the incidence of CRC is lower in Asian than in Western countries, recent studies have shown increasing rates of CRC in developing countries including Iran.[4,6,7]

Genetic syndromes such as Lynch and familial adenomatous polyposis, together with positive family history constituting up to 30% of CRC susceptibility; however, common genetic variants present in individuals’ genome considered as the main genetic risk of CRC.[8] Recent genome-wide association studies determined multiple CRC-related single nucleotide polymorphisms (SNPs).[9-13]

MicroRNAs (miRNAs) constitute a novel discovered a class of small noncoding RNAs (about 22 nucleotides long) that play important roles in posttranscriptional gene regulation by binding to the 3′-untranslated regions (3′-UTRs) of specific mRNAs resulting in mRNA cleavage or translational repression.[4,14-16] SNPs located in miRNAs-binding sites contribute to the disruption of miRNA recognition elements (MREs) or create new sites which eventually can lead to gene expression dysregulation and disease susceptibility.[2,17,18]

The previous studies have suggested that the SNP rs16917496 (CC genotype), which is located within the miR-502 binding site of SET8 (also known as SETD8 and PR-SET7) 3′UTR, is correlated with reduced SET8 protein levels.[19-21] SET8 is known as a histone H4-Lys-20 specific methyltransferase and its recognition sequence of miR-502 within 3′-UTR of SET8.
has been recently reported that modulates p53 activity by particularly monomethylation p53 at lysine 382 (p53K382me1).

\[22,23\] SET8 is involved in vital cellular processes such as transcriptional regulation, genomic stability, heterochromatin formation, DNA replication, and cell cycle arrest.\[24-26\] Therefore, it is proposed that deregulation of SET8 may affect individual’s cancer risk. Considering regulatory effects of SET8 on p53 activity, in conditions such as DNA damage, with decreased expression of SET8, the proapoptotic and checkpoint activation functions of p53 would be enhanced consequently. However, CC genotype related to SNP rs16917496 could disrupt this regulation.\[22,27,28\]

According to the available reports, this polymorphism is associated with breast, ovarian, lymphoblastic leukemia and cervical cancer susceptibility, and clinical outcome of hepatocellular carcinoma (HCC).\[19,20,27-30\] However, CRC susceptibility in relation to SNP rs16917496 has not worked out yet.

Here, we report the results of our case–control study on the correlation of SNP rs16917496 with sporadic CRC in the Iranian population. In addition, other factors such as gender, age, body mass index (BMI), physical activity, smoking status, and nonsteroidal anti-inflammatory drugs (NSAIDs) consumption were also analyzed in the present study.

**Materials and Methods**

**Study population**

For this case–control study, a total of 340 individuals were selected among patients referred to the colonoscopy centers of the Isfahan University of Medical Sciences Hospitals, from mid-2013 to mid-2015. One hundred and seventy CRC cases were selected on the basis of diagnosis by colonoscopy and then followed their pathology report for the ultimate confirmation of their colonoscopy-based CRC diagnosis. Equal numbers of individuals who underwent colonoscopy procedure but diagnosed with no malignancy or other adverse colonic diseases were constituted our normal control group. All participants filled and signed an informed consent form according to the Isfahan University of Medical Sciences Ethical Committee instructions. To eliminate any established genetic risk factor for CRC, patients were individually interviewed to select cases with sporadic CRC without positive history of familial cancers. In addition, all participants were asked to fill up a structured questionnaire to register the parameters known to influence the CRC susceptibility risk including gender, age, BMI, physical activity, smoking status, and NSAIDs consumption (for NSAIDs consumption any person who use aspirin or nonaspirin NSAIDs occasional or daily for long term [≥5 years] were considered as an NSAIDs consumer or regular consumer).

**Single nucleotide polymorphism selection**

Figure 1 summarizes the course of SNP selection in SET8 gene. As it is evident, SNPs located at 3'UTR of the target gene were considered for this study. Minor allele frequency higher than 0.05 and being located at MREs of miRNAs were the other two criteria considered for SNP selection. SNP rs16917496 was well fitted into our criteria and hence selected for this study.

**Single nucleotide polymorphism genotyping**

Genomic DNA was extracted from peripheral blood using PrimePrep Genomic DNA Isolation Kit (Genet Bio, Korea). The quality and quantity of the extracted DNA were assessed by agarose gel electrophoresis and spectrophotometry. SNP genotyping performed by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) method. PCR product size, restriction enzyme used, fragments subsequent to digestion, and primer sequences used are given in Table 1. PCR was performed in a total volume of 25 μl, including 2.5 μl of 10 × PCR buffer, 1 mM MgCl₂, 200 μM each dNTPs, 0.2 μM each primer, 100 ng of genomic DNA, and 1.5 U of Taq DNA polymerase (provided from CinnaGen company, Iran). Thermal cycling condition included an initial denaturation of 94°C for 5 min and then 34 cycles of 94°C, 62°C, and 72°C, all for 30 s and a final extension at 72°C for 5 min. RFLP was carried out using SwaI restriction endonuclease enzyme (provided from Thermo Fisher Scientific Company, USA). Subsequently, digested products were differentiated.
on agarose gel to determine the genotype of each sample. Fragment sizes of 390 and 360 bp were indicative of TT homozygous genotype; a single 750 bp fragment displayed the presence of CC homozygous and three fragments of 390, 360, and 750 bp indicated TC genotype [Figure 2]. Confirmation of the RFLP-based genotyping was done by direct sequencing of randomly selected samples with different genotypes (10% of samples).

**Statistics**

All statistical analyses were performed using SPSS version 22 (SPSS, Chicago, IL, USA). Genotype frequencies in cases and controls were tested for Hardy–Weinberg equilibrium using Chi-square test. Associations between SET8 (rs16917496) polymorphism with susceptibility to sporadic CRC were examined by logistic regression analysis. Age and sex were used as confounders. Correlation between this polymorphism and CRC was evaluated using odds ratios (ORs) and 95% confidence intervals (CIs). The significance level was set at $P < 0.05$. The difference in demographic and lifestyle characteristics distribution such as age, gender, smoking status, BMI, and NSAIDs consumption assessed by Pearson Chi-square test for categorical variables and $t$-test for continuous variables. Mann–Whitney test were used to compare physical activity between CRC and control groups.

**Results**

**Demographic and lifestyle characteristics**

Our study was intended to evaluate the relationship between rs16917496 SNP located at 3'UTR of SET8 gene and CRC susceptibility in a subset of the Iranian population. In this case–control study, 170 confirmed CRC patients (90 males and 80 females) and 170 (66 males and 104 females) unaffected individuals were selected among those referred to colonoscopy units of our University hospitals. Mean age in case and control groups was $57.72 \pm 12.29$ and $49.05 \pm 13.79$ years, respectively. Adjusted data (by age and gender) demonstrated that there was a considerable difference between case and control groups regarding physical activity ($P < 0.001$). Moreover, we found that healthy subjects in control group had more NSAIDs consumption compared with CRC patients group ($P < 0.004$). However, there were no statistically significant differences between patients and controls in terms of BMI and smoking status ($P = 0.19$ and $P = 0.23$, respectively). Characteristics of the study population are summarized in Table 2.

**Genotype and allele distribution**

The genotype distribution of rs16917496 polymorphism in cases and controls was in agreement with Hardy–Weinberg equilibrium. The frequency of TT, CC, and TC genotype in control group was 40.6%, 17.6%, and 41.8%, respectively and this frequency in case group was 34.1%, 18.8%, and 47.1%, respectively. In addition, the frequencies of T and C alleles were 61.5% and 38.5% in controls, and 57.6% and 42.4% in cases, respectively. Statistical analysis revealed that no significant correlation exists between SNP rs16917496 and risk of CRC in Iranian population ($P$ value for genotype and allele distribution, respectively: 0.32 and 0.15) [Table 3]. Because C is the ancestral allele for this SNP, we compared other two genotypes against CC

| SNP ID     | Primer sequence                  | PCR product length (bp) | Restriction enzyme (Ta°C) | RFLP fragment size (bp) |
|------------|----------------------------------|-------------------------|--------------------------|-------------------------|
| rs16917496 | F: CCCAGGAATCAAGGATAGGGC         | 750                     | SWAI (30°C)              | TT: 390+360             |
|            | R: GAGGTGGAGGTTGGAGGTTG          |                         |                          | CC: 750                 |
|            |                                  |                         |                          | CT: 750+390+360         |

PCR: Polymerase chain reaction, RFLP: Restriction fragment length polymorphism, SNP: Single nucleotide polymorphism

Figure 2: Agarose gel electrophoresis analysis of restriction fragment length polymorphism-polymerase chain reaction products. Lines 1 and 8 are TT genotype, 2–7 and 10 are TC genotype, and 9 is CC genotype
Here, we report an association between SNP rs16917496 genotypes and sporadic CRC risk in Iranian population. SET8 encodes a histone H4 methyltransferase that is essential for various biological processes, such as transcriptional regulation, heterochromatin formation, DNA replication, cell-cycle arrest, and maintenance of genome integrity. SET8 methyltransferase activity modifies p53 protein function by monomethylation of p53 at lysine 382. In cellular stresses including DNA damage, p53-dependent cell cycle arrest or apoptosis activated for which depletion of cellular SET8 is an essential prerequisite as seen that p53K382me1 levels decrease with DNA damage. Given, the importance of SET8-P53 interaction in cell-cycle and apoptosis control, deregulation of SET8 under influence of SNP rs16917496 T>C within the miR-502 binding site in the 3’-UTR of the SET8 gene would adversely affect this vital axis in the cell.

The previous studies demonstrated that SNP rs16917496 (CC genotype and C allele) which is located within the SET8 miR-502 binding site in its 3’UTR, is correlated with reduced SET8 protein levels. Here, we report an association between SNP rs16917496 genotypes and sporadic CRC risk in Iranian population. SET8 encodes a histone H4 methyltransferase that is essential for various biological processes, such as transcriptional regulation, heterochromatin formation, DNA replication, cell-cycle arrest, and maintenance of genome integrity. SET8 methyltransferase activity modifies p53 protein function by monomethylation of p53 at lysine 382. In cellular stresses including DNA damage, p53-dependent cell cycle arrest or apoptosis activated for which depletion of cellular SET8 is an essential prerequisite as seen that p53K382me1 levels decrease with DNA damage. Given, the importance of SET8-P53 interaction in cell-cycle and apoptosis control, deregulation of SET8 under influence of SNP rs16917496 T>C within the miR-502 binding site in the 3’-UTR of the SET8 gene would adversely affect this vital axis in the cell. Our data are indicative of no significant association between allele and genotype frequency of this polymorphism and sporadic CRC risk in Iranian population. Reviewing previous studies regarding rs16917496 T>C shows contradictory results. In Chinese population, CC and CC+CT genotypes are reported to be associated with longer survival of HCC, nonsmall-cell lung carcinoma (NSCLC) (CC genotype), and SCLC (CC+CT genotype) patients. In addition, the other study in China indicated that CC genotype is associated with decreased risk of epithelial ovarian cancer and TT genotype is associated with increased risk of NSCLC. On the other hand, a few studies revealed that CC genotype was associated with increased risk of breast cancer among the premenopausal women in the Chinese population, and CT+TT genotype is associated with decreased risk of acute lymphoblastic leukemia in comparison with CC genotype in the Iranian population. Despite these inconsistent results, all analysis on SET8 expression indicated that C allele is associated with reduced SET8 protein level.
Our results revealed no correlation of this polymorphism with sporadic CRC in Iranian population. In stratification analysis, neither CC genotype nor combined TT and TC genotypes, did not alter the risk of CRC among individuals under and over 55 years ($P = 0.40$ and $P = 0.38$, respectively), irregular and regular NSAIDs consumers ($P = 0.39$ and $P = 0.42$, respectively), and BMI under 25 and over 25 kg/m$^2$ ($P = 0.27$ and $P = 0.22$, respectively). However, in the subgroup of smokers, combined TT+TC genotype was associated with increased risk of sporadic CRC. However, there may be other players taking part in SET8 regulatory pathway. For instance, we can mention SCF (β-TRCP) which is marking SET8 for ubiquitination and degradation in a casein kinase I-mediated manner.$^{37}$

Our finding regarding rs16917496 T allele correlation with CRC in smokers compared to nonsmokers, would be an additional proof of cigarette health hazard. However, our findings regarding genotype-smoking interaction highlight this important point that certain genotypes are more readily accommodate adverse health effect of smoking compared to the other genotypes. Finally, it is the first evaluation of SNP rs16917496 located within the SET8 miR-502 binding site in relation with CRC. However, the reasons for the contradictory results may be owing to differences in ethnic background, and the different forms of CRC entities (familial vs. sporadic) and low-sample size that needs to more investigation in other population and more sample size.

**Acknowledgment**

We would like to appreciate financial support provided by Isfahan University of Medical Sciences.

**Financial support and sponsorship**

The study was supported by Isfahan University of Medical Sciences.

**Conflicts of interest**

There are no conflicts of interest.

**References**

1. Huang SH, Chao Y, Wu YY, Luo JC, Kao CH, Yen SH, et al. Concurrence of UGT1A polymorphism and end-stage renal disease leads to severe toxicities of irinotecan in a patient with metastatic colon cancer. Tumori 2011;97:243-7.
2. Chae YS, Kim JG, Lee SJ, Kang BW, Lee YJ, Park JY, et al. A miR-146a polymorphism (rs2910164) predicts risk of and survival from colorectal cancer. Anticancer Res 2013;33:3233-9.

3. Saito R, Suzuki H, Yamada T, Endo S, Moriwaki T, Ueno T, et al. Predicting skin toxicity according to EGFR polymorphisms in patients with colorectal cancer receiving antibody against EGFR. Anticancer Res 2013;33:4995-8.

4. Bhatumuk P, Gopalakrishnan C, Kamaraj B, Purohit R. Single nucleotide polymorphisms in microRNA binding sites: Implications in colorectal cancer. ScientificWorldJournal 2014;2014:547154.

5. Feik E, Baierl A, Hieber B, Führlinger G, Pentz A, Stättner S, et al. Association of IGF1 and IGFFBP3 polymorphisms with colorectal polyps and colorectal cancer risk. Cancer Causes Control 2010;21:91-7.

6. Kolahdoozan S, Sadjadi A, Radmard AR, Khademi H. Five common cancers in Iran. Arch Iran Med 2010;13:143-6.

7. Hosseini SV, Izadpanah A, Yarmohammadi H. Epidemiological changes in colorectal cancer in Shiraz, Iran: 1980-2000. ANZ J Surg 2004;74:547-9.

8. Chang WS, Yueh TC, Tsai CW, Ji HX, Wu CN, Wang SC, et al. Contribution of DNA repair xeroderma pigmentosum group D Genotypes to colorectal cancer risk in Taiwan. Anticancer Res 2016;36:1657-63.

9. Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet 2007;39:984-8.

10. Tenesa A, Farrington SM, Prendergast JG, Porteous ME, Walker M, Haq N, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nat Genet 2008;40:631-7.

11. Zhang K, Civan J, Mukherjee S, Patel F, Yang H. Genetic variations in colorectal cancer risk and clinical outcome. World J Gastroenterol 2014;20:4167-77.

12. Daraei A, Salehi R, Salehi M, Emami MH, Janghorbani M, Mohamadhashem F, et al. Effect of rs6983267 polymorphism in the 8q24 region and rs4444903 polymorphism in EGF gene on the risk of sporadic colorectal cancer in Iranian population. Med Oncol 2012;29:1044-9.

13. Wu CZ, Ni XJ, Zheng SL, Yang YR, Xia P, Zeng YJ, et al. Prediction of microRNA target sites of 37 selected cancer-related genes and potential implications for cancer therapy. Nucleic Acids Res 2007;35:4535-41.

14. Walker M, Haq N, et al. The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells. Nat Cell Biol 2010;12:1086-93.

15. Julien E, Herr W. A switch in mitotic histone H4 lysine 20 methylation status is linked to M phase defects upon loss of HCF-1. Mol Cell 2004;14:713-25.

16. Tardat M, Brustel J, Kirsh O, Lefebvre C, Callanan M, Sardet C, et al. The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells. Nat Cell Biol 2010;12:1086-93.

17. Jørgensen S, Elvers I, Trelle MB, Menzel T, Eskildsen M, Jensen ON, et al. The histone methyltransferase SET8 is required for S-phase progression. J Cell Biol 2007;179:1337-45.

18. Xu J, Yin Z, Gao W, Liu L, Yin Y, Liu P, et al. Genetic variation in a microRNA-502 binding site in SET8 gene confers clinical outcome of non-small cell lung cancer in a Chinese population. PLoS One 2013;8:e77024.

19. Hashemi M, Sheyhani-Nasab M, Naderi M, Roodbari F, Taheri M. Association of functional polymorphism at the miR-502-binding site in the 3'-untranslated region of the SETD8 gene with risk of childhood acute lymphoblastic leukemia, a preliminary report. Tumour Biol 2014;35:10375-9.

20. Guo Z, Wu C, Wang X, Wang C, Zhang R, Shan B, et al. Polymorphism at the miR-502 binding site in the 3'-untranslated region of the histone methyltransferase SET8 is associated with hepatocellular carcinoma outcome. Int J Cancer 2012;131:1318-22.

21. Diao L, Su H, Wei G, Li T, Gao Y, Zhao G, et al. Prognostic value of microRNA 502 binding site SNP in the 3'-untranslated region of the SET8 gene in patients with non-Hodgkin’s lymphoma. Tumori 2014;100:533-8.

22. Shi X, Kachirskaya I, Yamauchi H, West LE, Wen H, Wang EW, et al. Modulation of p53 function by SET8-mediated methyltransferation at lysine 382. Mol Cell 2007;27:636-46.

23. Fang J, Feng Q, Ketel CS, Wang H, Cao R, Xia L, et al. Purification and functional characterization of SET8, a nucleosomal histone H4 lysine 20-specific methyltransferase. Curr Biol 2002;12:1086-99.

24. Diao L, Su H, Wei G, Li T, Gao Y, Zhao G, et al. Prognostic value of microRNA 502 binding site SNP in the 3'-untranslated region of the SET8 gene in patients with non-Hodgkin’s lymphoma. Tumori 2014;100:533-8.
associated with silent chromatin. Mol Cell 2002;9:1201-13.
35. Ding C, Li R, Peng J, Li S, Guo Z. A polymorphism at the miR-502 binding site in the 3′ untranslated region of the SET8 gene is associated with the outcome of small-cell lung cancer. Exp Ther Med 2012;3:689-92.
36. Wang C, Guo Z, Wu C, Li Y, Kang S. A polymorphism at the miR-502 binding site in the 3′ untranslated region of the SET8 gene is associated with the risk of epithelial ovarian cancer. Cancer Genet 2012;205:373-6.
37. Wang Z, Dai X, Zhong J, Inuzuka H, Wan L, Li X, et al. SCF (ß-TRCP) promotes cell growth by targeting PR-Set7/Set8 for degradation. Nat Commun 2015;6:10185.