Original Article

Mutant Allele of CD44 (rs8193C>T) and Pum2 Regulatory Element as A Prognosis Factor of Prostate Neoplasms: A Case-Control and In Silico Studies

Mohammadkazem Heydari, M.Sc.[1], Abasalt Hosseinzadeh Colagar, Ph.D.[1]*, Emadoddin Moudi, Ph.D.[2]

1. Department of Molecular and Cell Biology, Faculty of Science, University of Mazandaran, Babolsar, Iran
2. Department of Urology, Babol University of Medical Sciences, Babol, Iran

Corresponding Address: P.O.Box: 47416-95447, Department of Molecular and Cell Biology, Faculty of Science, University of Mazandaran, Babolsar, Iran
Email: ahcolagar@umz.ac.ir

Received: 13/February/2022, Accepted: 08/June/2022

Abstract

Objective: Expression of CD44 variant 6 (CD44v6) as a homing-associated cell adhesion molecule (HCAM), has proved to change most cancer cells. Aim of the study is the effect of mutant allele of *CD44* (rs8193C>T) and Pum2 regulatory element as a prognosis factor of prostate neoplasms: a case-control and in silico studies in the Mazandaran province-Iran.

Materials and Methods: In a case-control study, CD44-rs8193C>T genotyping of the 420 prostate neoplasms (210 benign prostatic hyperplasia (BPH) patients and 210 prostate cancer patients) and 150 healthy samples are performed by the touchdown polymerase chain reaction with confronting two-pair primers (PCR-CTPP) method. The T mutant allele effects on the mRNA structure and cell pathways were also investigated in silico methods.

Results: Our results showed that the increase of T mutant allele frequency was significantly associated with BPH compared with prostate cancer. Furthermore, results showed TT genotype was significantly associated with BPH [odds ratio (OR)=0.572 and P=0.015], and also influenced the CD44v6 transcript secondary structure, miRNA binding, and regulatory element-binding site for Pum2 protein. Attachment of Pum2 to standard CD44 transcript may lead to transcript isoform-switching and shift-expression to a variety of CD44 isoforms, which can trigger some of the cell signaling pathways, such as Nanog-Stat, PKC-Nanog, and PKC-Twist.

Conclusion: Based on this, the presence of the T mutant allele of CD44 (rs8193C>T) in the populations may create a regulatory element-binding site for Pum2. So, it could be known as a prognosis factor and prediction of prostate neoplasms. However, more comprehensive studies in different populations (with various ethnicities and large population sizes), and also CD44v6 gene expression studies in protein and transcript levels are required to confirm our data.

Keywords: CD44, Neoplasm, Prostate, Pum2

Cell Journal(Yakhteh), Vol 24, No 12, December 2022, Pages: 723-731

Citation: **Heydari M, Hosseinzadeh Colagar A, Moudi E. Mutant allele of CD44 (rs8193C>T) and Pum2 rgulatory element as a prognosis factor of prostate neoplasms: a case-control and in silico studies. Cell J. 2022; 24(12): 723-731. doi: 10.22074/cellj.2022.8468.**

Introduction

The prostate neoplasm or tumor refers to a group of cells with unregulated growth and cell proliferation. There are two types of prostate tumors: benign prostatic hyperplasia (BPH) and malignant tumor, malignant tumor generally known as prostate cancer (1). Prostate cancer is known as the second most frequent cancer in men and the worldwide fifth leading cause of death. There are several risk factors for prostate cancer including age, ethnicity, family history, genetic disorders, or genetic risk factors (2). The gene risk factors occur as a chromosome and nucleotide-level genomic instability. The chromosome instability may happen due to the change of number (aneuploidy), structure (e.g., translocation, inversion, or duplication), or both. However, nucleotide-level genomic instability includes single nucleotide variations (SNVs) and small insertion/deletions (3). Single-nucleotide polymorphisms (SNPs) are one of the SNVs by redundancy upper than 1%. They happen in a variety of genes, correlated to prostate neoplasm e.g., *HPC1*, *HPC2* (4), *MSMB*, *KLK2-3* (5), *HNF1B* (6), and *CD44* (7).

CD44 is an integral cell surface protein that has been referred to as a homing-associated cell adhesion molecule (HCAM), phagocytic glycoprotein-1 (Pgp-1), Hermes antigen, lymphocyte homing receptor, extracellular matrix receptor III (ECMRIII) (8), and HUTCH-1. CD44 is a mouse IgG2a mAb derived from the fusion of NS-1 myeloma cells to spleen cells from a BALB/c mouse immunized with whole macaque peripheral blood mononuclear cell/PBMC (9). In humans, its protein-coding gene is located on chromosome 11p13 with about 443500 base pairs and 20 exons. Its largest variant CDS has a 5437 bp with 19 exons because exon 19 is normally spliced out (10).

CD44 protein contains four domains: two N-terminal extracellular regions (globular- and stem), a helical transmembrane domain, and a cytoplasmic domain (10). The globular- extracellular region, which is glycosylated and acts as a ligand-binding site, is encoded by exons 1 to 5, stem- the extracellular region is encoded by exons 6 to 17, the helical domain is encoded by exon 18, and cytoplasmic domain (C-terminal) is encoded by exon 20 (8). Alternative splicing of hnRNA-CD44 transcripts led to CD44 different isoforms (CD44v), which named CD44v1 to CD44v10. These CD44vs may have individual or incorporations

sequences encoded by exons 6-15 (also named CD44v1-CD44v10), inserted in the stem in the extracellular region (10). When all variable exons (v1-v10 in humans) are spliced out it is called standard CD44 (CD44s).

Some researchers commented that CD44 isoforms may have interplay with a variety of cell signaling pathways so, it is considered to be a signaling hub and interact with several ligands. Some of these ligands are protein receptors including EGF, TGF-β (11), etc., and extracellular martial including serglycin, osteopontin, collagen (12), etc. These interactions can lead to proliferation, adhesion, migration, and invasion of cells (13). On the other hand, researchers emphasized that *CD44* gene variants can lead to various diseases (14, 15).

Previous researchers reported that CD44 receptors, as CAMs, play an important role in metastasis and homing of cells (16). Also, their expressions inhibit tumor growth (17), so their gene modification is associated with several cancers (15). On the other hand, some researchers reported that two isoforms of CD44 (CD44s and CD44v6) are very important when they are expressed by alternative splicing of a single mRNA in the prostate neoplasm (18). These researchers showed that expression of CD44s in prostate adenocarcinoma is lost and may predict a poor prognosis independent of stage and grade. Heider et al. (19) also characterized a high-affinity monoclonal antibody specific for CD44v6, which is a candidate for immunotherapy of squamous cell carcinomas. Certain CD44 isoforms, especially CD44v6 have been implicated in tumorigenesis, tumor cell invasion, and metastasis (20).

We hypothesize that single nucleotide mutations in populations alter transcription factors or create new locations for regulatory factors on the gene transcripts, which are the shift expression of these protein isoforms. Previous studies showed the correlation of the CD44 SNPs (21) to several cancers including breast- (22), prostate- (7), and bladder- (23) cancers. For example, rs187115T>C of CD44 is associated with non-small cell lung cancer/ NSCLC risk (24). Also, rs353639T>G of CD44 may have a risk for bladder cancer (23), and rs8193C>T, which is located in the 3′UTR area of all variants except for CD44v8, is statistically associated with the risk of gastric cancer (25).

In this study, the association of the CD44-rs8193C>T mutant allele with prostate neoplasms (BPH and prostate cancers) (North of Iran), as well as its probable effects on CD44v6 structure and cell signaling pathways have been studied in Mazandaran province.

Materials and Methods

Samples collection

The intravenous blood of 570 samples including two case groups of BPH (210 patients) and prostate cancers (210 patients) without any cancer treatments such as prostatectomy, radiotherapy, and chemotherapy and 150 healthy men, without any history of prostate cancer, as a control group was included in this case-control study. The patients were histologically verified for BPH and prostate cancer, and control groups without any history of prostate cancer. Five milliliters of the case and control blood from the same geographical origin (Mazandaran, Iran) were collected into tubes containing $EDTA_{Na2}$ as an anticoagulant. All samples were prepared from some Mazandaran hospitals (Nimeh-e-Shaban, Sari; Babol Clinic and Rhohani Hospital of Babol, Iran) from September 2018 to March 2020.

This study is approved by the Ethics Committees of Mazandaran University of Medical Science (IR.UMZ.REC.1400.001) and all subjects signed an informed consent form before entering the study.

Genotyping by touchdown PCR-CTPP

Genomic DNA is extracted by the phenol-chloroform DNA extraction method from white blood cells (26). The CD44 genotypes were detected with the PCR-confronting two-pair primers (PCR-CTPP) method, as DNA fragments of the SNP (rs8198C>T) were amplified with four primers (Table 1). Touchdown PCR-CTPP fragment lengths are shown in Figure 1. Touchdown PCR program was used for amplification of DNA fragments, under Table 1 conditions in 25 µl of PCR total volume, by DNA thermal cycler (Master Cycler Gradient, Eppendorf Co., Germany).

Table 1: PCR conditions and oligomers used as primers

Primers names	Oligomers (5′-3′)*	PCR conditions	Cycles and thermal conditions	Twin primers: PCR products (bp)
C8193T-OF	TGTATTCCTGATCGCCAACCTTTC	0.25 U, Taq polymerase; 0.2 mM, mix dNTP;0.35 µM each of forward and reverse primers; 4 mM, $MgCl_2$; 7.5%, DMSO; 0.1 mg/ml, BSA; 2.5 µl of 10X AMS buffer (750 mM, Tris-Cl, pH=8.8; 200 mM, $(NH4)_2SO_4$; 0.1% Tween 20); and ~30 ng of t-DNA	4 minutes at 94˚C; 2 cycles (45 secods at 94˚C, 30 seconds at 64˚C, 30 seconds 72 ˚C); 2 cycles (45 seconds at 94˚C, 30 seconds at 63˚C, 30 seconds 72 ˚C); 2 cycles (45 seconds at 94˚C, 30 seconds at 62˚C, 30 seconds 72 ˚C); 2 cycles (45 seconds at 94˚C, 30 seconds at 61˚C, 30 seconds 72 ˚C); 25 cycles (45 seconds at 94˚C, 30 seconds at 60˚C, 30 seconds 72 ˚C); Final exten:10 minutes at 68-72˚C	*C8193T-OF / -OR*: 533
C8193T-OR	ATGAGATTTGGCTGAGTGGGTC			*C8193T-IF / -OR*: 427
C8193T-IF	GCCTAATCCCTGGGCATTGT			*C8193T-OF / -IR*: 147
C8193T-IR	CCCCAACCTCAGTGGAAAG			

PCR; Polymerase chain reaction, * Oligomers; Synthesized by CinnaClon Co, Iran, bp; Base pair, t-DNA; Template DNA, OF; Outer forward, IF; Inter forward, OR; Outer revers, and IR; Inter revers.

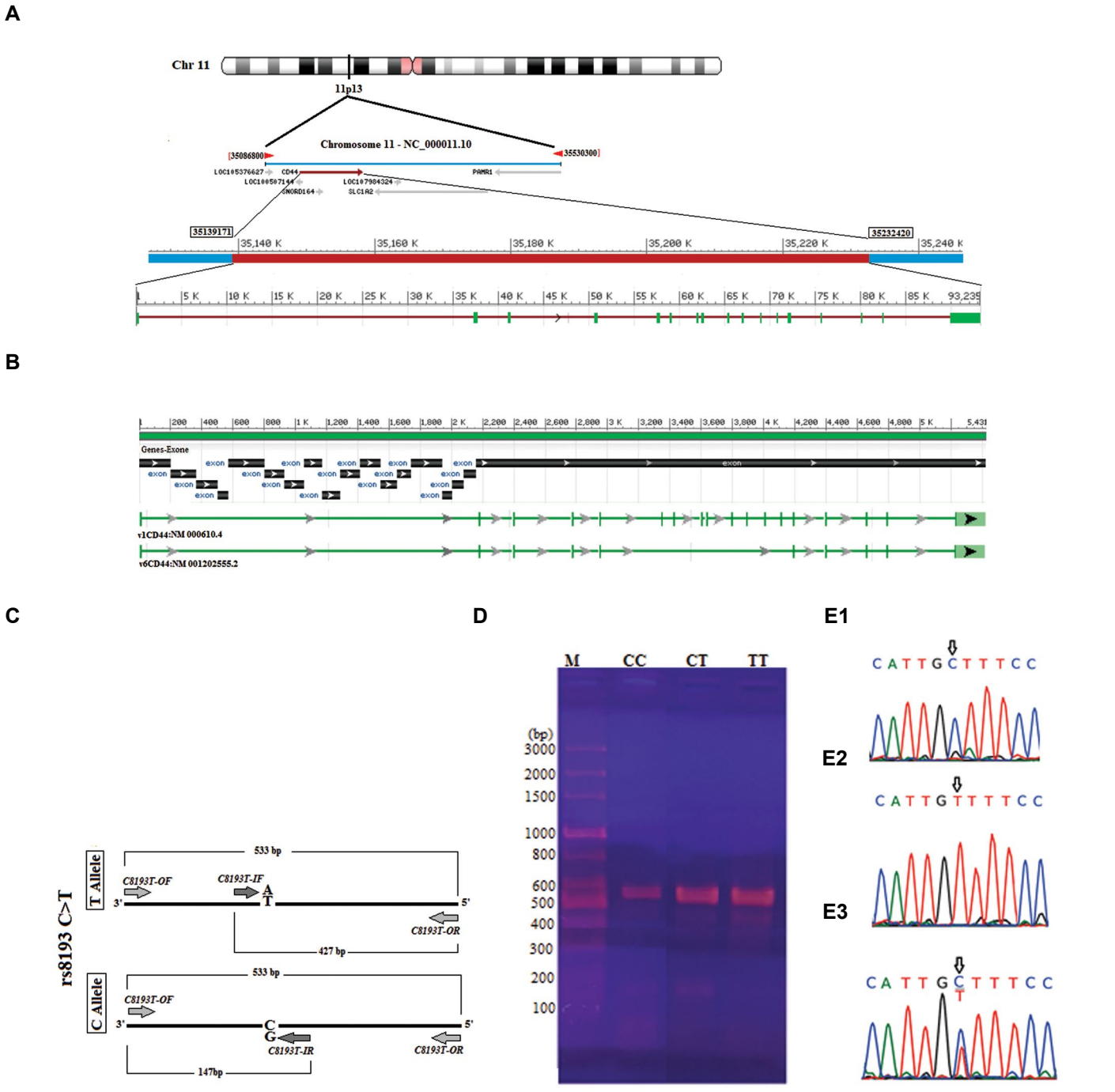

Fig.1: Schematic drawing of the *CD44* gene region on chromosome 11, and Touchdown PCR-CTPP mapping. **A.** Map of *CD44* gene with 93249 bp, retrieved from NCBI database # NC:000011.10. **B.** mRNA transcript of CD44v1 with 18 exons and 5431 bp, as the longest transcript, and CD44v6 transcript with 10 exons and 4492 bp, which studied in this paper. **C.** Touchdown PCR-CTPP diagrammatic map and its products length, that performed by four primers for diagnostic T allele and C allele of the rs8193C>T. **D.** PCR product pattern in the 1.5% agarose gel electrophoresis, which stained by 0.1 µg/ml ethidium bromide. CC, TC, and TT lanes are genotypes, M; Mid-range DNA ladder (Jena Bioscience JmbH Co. Germany), **E1.**, **E2.** and **E3.** Electropherogram results of PCR direct sequence of CC, TT, and TC genotypes, respectively. PCR-CTPP; Touchdown polymerase chain reaction with confronting two-pair primers.

All PCR reactants and chemical materials were purchased from SinnaClon Co (Iran) and Merck Co. (Germany), respectively. PCR fragments were electrophoresed in 1.5% agarose gel, stained in the 1 µg/ml of ethidium bromide solution, and observed by UV-transilluminator (ProteinSimple Red SA-01587, Co, USA). Then, at least three PCR product samples from any genotypes were sequenced by Bioneer Co. (Korea) for confirmation of PCR-CTPP results. Subsequently, electropherograms of the sequences were analyzed by Chromas ver 2.0 software.

Data sources for in silico analysis

The bioinformatics analysis used for SNPs selection is based on the information of the NCBI (https://www.ncbi.

nlm.nih.gov) data bank. To investigate and influences of rs8193C>T on the stability of the CD44 (mRNA) RNAsnp web server was used (https://rth.dk/resources/rnasnp). RegRNA2.0 (http://regrna2.mbc.nctu.edu.tw) is used for UTR motifs and their regulatory transcription factors. For the presence or absence of microRNAs (miRNAs), the miRNA SNP-v3 database (http://bioinfo.life.hust.edu.cn/miRNASNP) is used by targets gain/loss by SNP in miRNA seed modules, which provides the effect of SNPs in miRNA seed regions. In this module, at first miRNA wild sequence and SNP allele sequence were given. Then, two target prediction tools Targetscan (http://www.targetscan.org/vert_72/) and Miranda (http://www.microrna.org/microrna/home.do) web servers were used for the prediction of target sites, respectively. If one target shows in both SNP Targetscan (ST) and SNP Miranda (SM) web servers but not in either Wild Targescan (WT) or Wild Miranda (WM) web servers, it is called the miRNA gained one target gene. Study of signaling pathway performed by Kyoto encyclopedia of genes and genomes (KEGG) database (https://www.genome.jp/kegg/pathway.html). *CD44* gene expression, both on the RNA transcript and protein levels, in the various human tissues, was done by the human protein atlas (HPA) data bank (https://www.proteinatlas.org/).

Post-transcriptional UTR regulatory motif elements of CD44v6-rs8193C>T were analyzed by scanning for motifs (SFM) free web server (http://crispr.otago.ac.nz/). In this web server regulatory elements calculate with a matching score <= 0.175 per thousand bases analysis from the Transterm database (http://crispr.otago.ac.nz/newTTDB/) and protein binding sites from RNA binding protein database (RBPDB) (http://rbpdb.ccbr.utoronto.ca/index.php), which frequency matrices are available with an E-value <= 1.

Statistical analysis

Statistical genotype and allele differences and frequencies between normal and BPH; normal and cancer groups were analyzed by a Chi-squared ($\chi 2$) test so odds ratios (ORs) and 95% confidence interval (95% CI) were calculated for various alleles and genotypes. A two-tailed $P<0.05$ was considered statistically significant. All statistical analyses were performed by SPSS ver. 19 (SPSS Inc., IBM Corp Armonk, NY, USA).

Results

In silico analysis of CD44-SNPs

In the preliminary of the CD44-SNPs, we found that this gene contains 156 SNPs loci with allele frequency (MAF) ≥ 0.1 (https://www.ncbi.nlm.nih.gov). From these SNPs, 12 were intron-SNPs (rs7116432A>G, rs7110737T>A, rs379410T>C, rs3794109A>G, rs3794105G>A, rs353647C>G, rs353637A>T, rs353630G>A, rs353623G>A, rs353618T>C, rs353612G>A and rs112762C>T) which associated with clinical patients (https://www.ncbi.nlm.nih.gov/clinvar/), and five of them are 3'UTR-SNPs (rs7116432A>G, rs11607862C>T, rs7116739G>C, rs13347C>T and rs8193C>T). Three SNPs (rs7116432A>G, rs11607862C>T, rs7116739G>C) are in the 3'UTR region of variant 8 of CD44 (CD44v8), but other variants (CD44v1-CD44v10) were in intron regions. Furthermore, two SNPs (rs13347C>T and rs8193C>T), which are located in the 3'UTR area, were found in other variants (including CD44v6) except CD44v8. On the other hand, MAF of rs8193C>T (0.35 and 0.65 for C- and T- alleles, respectively), which have been reported in South Asian populations in the 1000 Genomes project, was selected in this study in Mazandaran province (North of Iran) populations.

Studying the effects of rs8193C>T on the stability of the CD44v6 transcript using the RNAsnp server showed that these SNPs loci have no significant effect on the stability of the CD44v6 transcript (P=0.2969). While there is a slight change in the structure of the CD44v6 transcript in both the mutant and wild forms in rs8193C>T, which is associated with a change in free energy from -109.30 to -111.30 (Fig.2A1-4).

Analysis of the post-transcriptional regulatory elements analysis of CD44v6 -rs8193C>T showed that the wild allele of this variant (C-allele) matches to P-element somatic inhibitor (Psi), embryonic lethal, abnormal vision- like 1 (ELAV1) and Y box binding protein 1 (YBX1) but T-mutant allele of CD44v6-rs8193C>T match to ELAV1, Pumilio homolog 2 (Pum2), and Psi (Fig.2B). The effects of rs8193C>T variants on affinity CD44v6 transcript with miRNAs were analyzed using the miRNA SNP-v3 database. As a result, the CD44v6 transcript gains four miRNAs and a loss of six miRNAs (Table 2). On the other hand, our study with the KEGG database showed that CD44 as a receptor can be bound to extracellular materials, especially hyaluronic acid (HA), and collaborated in at least six important pathways. These interactions are involved in angiogenesis, cell -growth, -survival, -differentiation, -invasion, and -migration. For example, it can result in "cell growth and survival" and "cell migration and invasion", by activating two downstream pathways with Nanog and Twist transcriptional factors.

The analysis of CD44 expression in two transcript and protein levels by the HPA databank showed that there is a significant difference between RNA transcript levels and protein levels in various tissues. So, it might be expressed in some tissues (e.g., stomach and liver) but its protein level is low or not enough for detection. Conversely, CD44-RNA may be

expressed at a lower level in some tissues (e.g., breast and prostate) but its protein expression is enough or at a high level for detection.

Genotyping

The rs8193C>T statistical analysis was performed between normal vs. BPH and normal vs cancer. These analyses showed that there is a significant difference in TT genotypes between normal and BPH groups (P=0.015) but not between normal and cancer groups (P=0.086). However, there is no significant difference in CT and CT+TT genotype between normal vs BPH and normal vs cancer groups. Also, results showed that the mutated allele (T- allele) with a frequency of 54% in the normal group, and 61.66% in the BPH group has a significant difference between normal and BPH (OR=0.730 and P=0.040). However, the T allele is not associated with prostate cancer (Table 3).

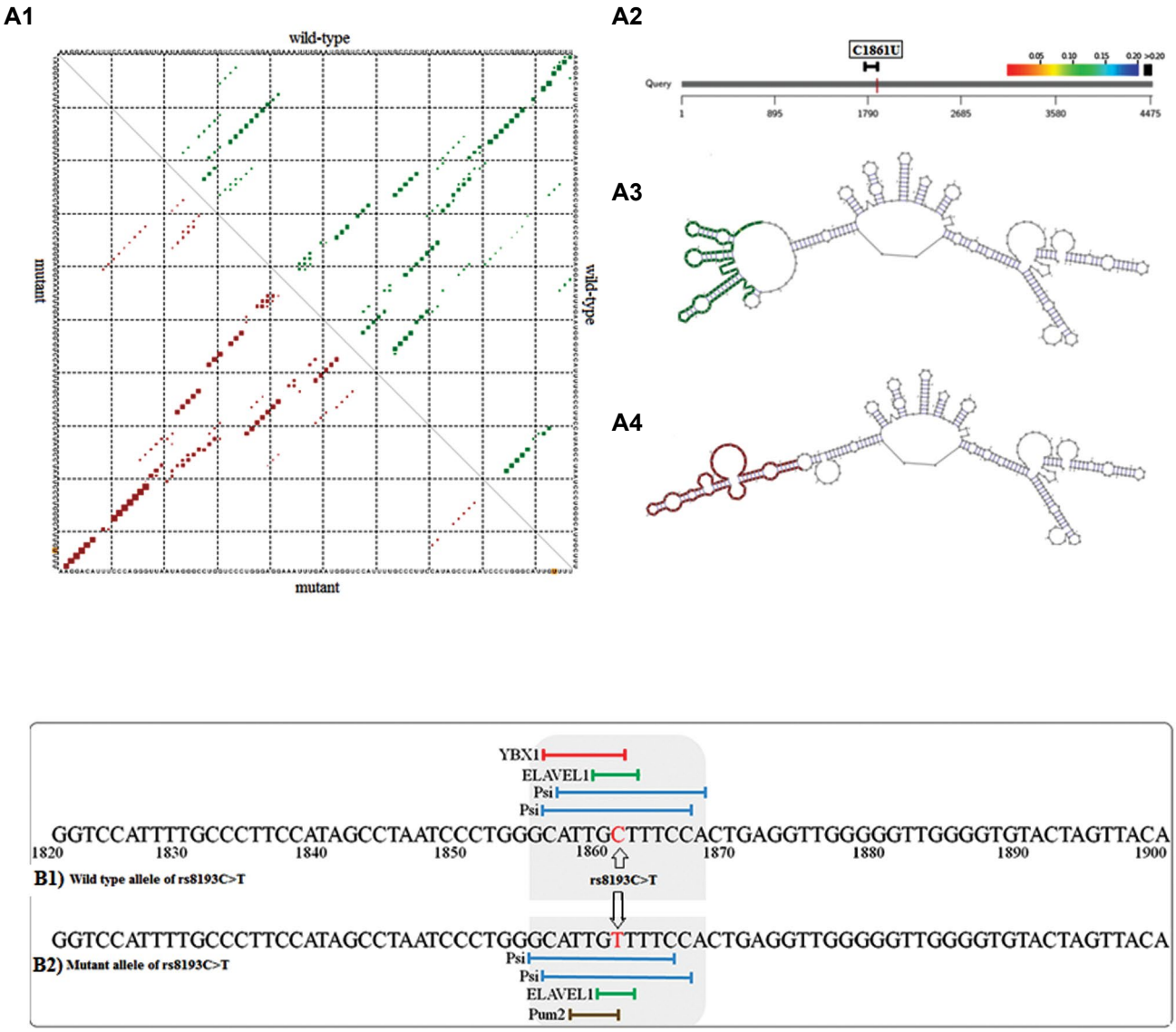

Fig.2: The result of in silico analysis. **A1-A4.** RNASNP analysis of rs8193C>T in the local region with maximum differences in wild-type and mutant mRNA. Base pair probabilities of the local region (1661-2061bp for rs8193C>T) were detected with maximum differences depicted. The upper and lower triangle of the matrix represents the base pair probabilities of wild-type and mutant sequences, respectively. **A1.** The mutated nucleotide is shown in yellow. **A2.** A graphic summary of the analysis. The SNP-affected region is colored in black since the P>0.2, which means not very significant change occurred in mRNA structure for rs8193C>T, and fewer than 0.2, which means very significant change occurred in mRNA structure for rs8198C>T with P value which is color direction. **A3.** The optimal secondary structure of global wild-type sequence (1661-2061 bp) depicted in green with the minimum free energy of -109.30 kcal/mol, and for rs8193C>T. **A4.** The optimal secondary structure of the global mutant sequence, 1661-2061bp, is shown in red with the minimum free energy equals to -111.30 kcal/mol for rs8193C>T; Retrieved from https://rth.dk/resources/rnasnp/results?jobid=579164-0212756279; 15 Feb 2020, 19:48 pm. **B.** Also, results of SFM webserver shown for **B1.** Wild type and **B2.** Mutant allele, respectively. SNP; single-nucleotide polymorphisms.

Table 2: In silico analysis rs8193C>T effects on the miRNA binding by miRNA SNP-v3 database

Gain/ Loss	miRNA	ΔG*duplex	ΔG binding	ΔG open	TargetScan score	AU content	Exact probability
Gain	hsa-miR-10523-5p	-9.70	-8.06	18.07	20.48	0.48	0.03
	hsa-miR-3148	-9.80	-10.11	18.15	21.97	0.47	0.05
	hsa-miR-570-3p	-8.90	-6.47	18.14	21.49	0.49	0.01
	hsa-miR-6124	-14.90	-15.13	18.15	23.47	0.47	0.05
Loss	hsa-miR-10527-5p	-16.20	-9.83	17.66	23.34	0.34	0.04
	hsa-miR-3908	-13.30	-11.47	17.66	21.38	0.38	0.01
	hsa-miR-3942-5p	-11.80	-6.69	17.66	20.42	0.42	0.02
	hsa-miR-4422	-11.20	-10.43	17.66	23.36	0.36	0.04
	hsa-miR-4703-5p	-11.60	-6.19	17.66	20.92	0.42	0.02
	hsa-miR-6835-3p	-10.30	-8.53	17.66	22.36	0.36	0.04

ΔG energy; kCal/mo, AU; Adenine and uracil, Gain; SNP in gene 3'UTR causes target gain, Loss; SNP in gene 3'UTR causes target loss, and SNP; Single-nucleotide polymorphisms.

Table 3: Genotype and allele frequencies of *CD44v6*-rs8193C>T were analyzed in case and control samples

Genotype	Normal (n=150)	BPH (n=210)	P value [OR (95% CI)]*	HWE	Cancer (n=210)	P value [OR (95% CI)]*	HWE
CC	30 (20)	36 (17.14)	Reference	0.395	42 (20)	Reference	0.465
CT	78 (52)	89 (42.38)	0.072 [1.473 (0.966-2.245)]	-	91 (43.33)	0.105 [1.417 (0.930-2.158)]	-
TT	42 (28)	85 (40.47)	0.015 [0.572 (0.364-0.897)]	-	77 (36.66)	0.086 [0.672 (0.427-1.057)]	-
CT+TT	120 (80)	174 (81.66)	0.490 [0.828 (0.484-1.417)]	-	168 (80)	0.912 [1.030 (0.611-1.736)]	-
C- allele	138 (46)	161 (38.33)	Reference	-	175 (41.66)	Reference	-
T- allele	162 (54)	259 (61.66)	0.040 [0.730 (0.540-0.985)]	-	245 (58.33)	0.248 [0.839 (0.622-1.130)]	-

Data are presented as n (%). OR; Odds ratio, CI; Confidence interval, HWE; P value for Hardy-Weinberg equilibrium, * OR (95% CI); OR (95% CI lower-upper).

SNP and isoform switching

Likely, the SNP of rs8193C>T led to the creation of a new regulatory element site, which is named the Pum2 protein binding site. This new regulatory element (Pum2) may cause translational inhibition and mRNA degradation. Also, it may lead to isoforms switch so, which results in overexpression of other variants or isoforms (Fig.3).

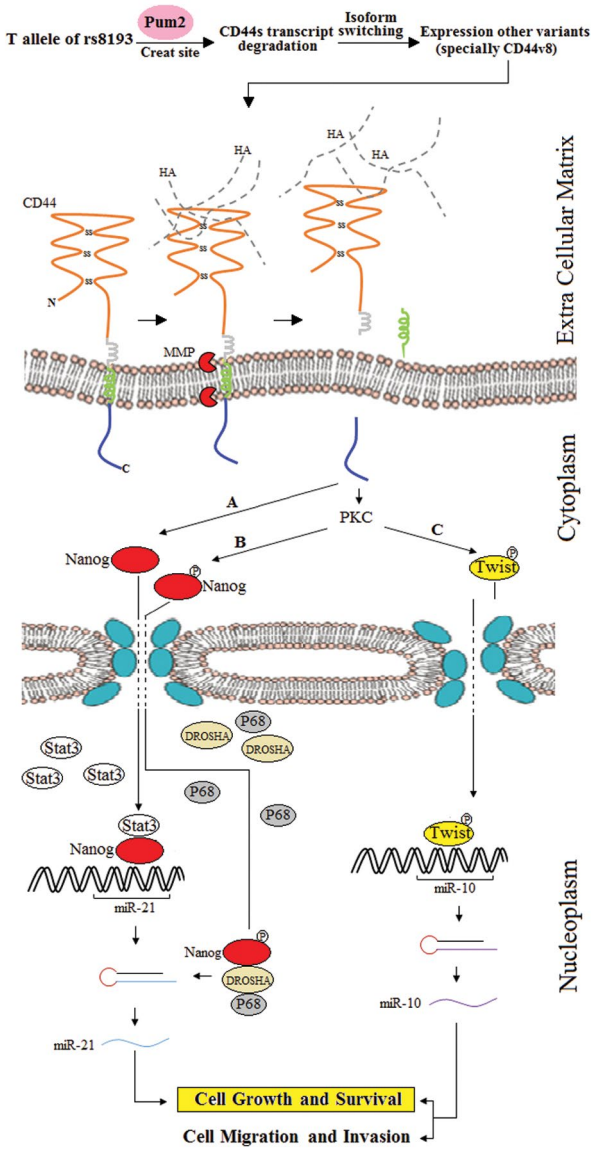

Fig.3: Effects of the CD44 rs8193C>T transition on its transcripts under mechanism of isoform switching. Pum2; Pumilio homolog 2, HA; Hyaluronic acid, CD44; Cluster of differentiation 44, MMP; Matrix metalloproteinases, PKC; Protein kinase C, Stat3; Signal transducer and activator of transcription 3, miR-21; MicroRNA-21, miR-10; MicroRNA-10, A; Nanog-Stat pathway, B; PKC-Nanog pathway, and C; PKC-Twist pathway.

Discussion

Prostate cancer is one of the most prevalent cancers among men. Risk factors of age, ethnicity, and family history could have an effect on the pathogenicity of prostate carcinogenesis (2). There are so many different genetic markers that have been reported to evaluate cancer susceptibility, including, *CD44* (7). All variant isoforms of CD44 are cell-surface glycoproteins, generated by alternative splicing and post-translational modifications (19). These CAMs are expressed in a variety of human tumors and play important roles in tumor progression, homing and metastasis (16), and inhibition of tumor growth (17). The study of its expression and gene modification has always been considered by researchers as a controversy. For instance, the correlation of SNPs of its different variants (CD44s and CD44v1-v10) (21), as a gene modification, was studied in a variety of cancers (22, 23). One of the SNPs of CD44v6, which was investigated by previous researchers is rs8193C>T. The effect of this SNP on the risk of gastric cancer (25), and colon cancer (27, 28) were studied before. The C/C genotype of CD44 rs8193C>T may be associated with higher transcriptional activity of this gene as predicted by F-SNP (29), but the analysis of the HPA data bank showed that the transcriptional activity and the level of protein produced were related to tissue types.

Previous studies showed that lncRNA-uc002kmd.1 known as a molecular decoy for miR211-3p and targets the degradation of CD44. Wu et al. (30) report that lncRNA-uc002kmd.1 mediated cell migration and proliferation in gastric cancer populations. But our study showed that there isn't any ncRNA, which matches rs8193C>T.

Gerger et al. (27) showed that the rs8193C>T variant represented a high-risk subgroup with time to tumor recurrence (TTR) in colon cancer patients, in the American population (California). Thus, patients carrying at least one T allele of CD44 rs8193C>T have longer TTR compared with patients with homozygous C/C (27). The study of Stremitzer et al. (28) in Southern California showed rs8193C>T associated with a higher radiological response rate so the C/C genotype is a lower rate compared with other genotypes, C/C 72%, any T 88%, P=0.033. Mokhtarian et al. (25) reported C allele of rs8193C>T is associated with the risk of malignancy, lymph node spread, and stage of gastric cancer, in the Iranian population (Esfahan province), while our result showed that the T allele of CD44 rs8193C>T was significantly associated with BPH with P=0.040 whereas prostate cancer with P=0.248. In addition, results showed that TT genotypes were significantly associated with BPH (P=0.015).

While Mokhtarian et al. (25) reported a correlation between rs8193C>T and gastric cancer, the results of Winder et al. (31) showed that there is no relation between rs8193C>T and gastric cancer. It seems that this difference is a result of different ethnicities.

Some researchers have reported that they did not observe any significant difference in rs8193C>T and susceptibility to breast cancer in Chinese women (32), gastric adenocarcinoma (31), and hepatocellular carcinoma (HCC) (33). However, this difference may be due to the different types of tissues and variety of the CD44 RNA and protein expression levels while our study also confirms that. Probably, rs8193C>T CD44v6 transition plays a different role in different ethnic populations and different cancer types.

The previous studies of CD44 isoform expression in the prostate neoplasm tissue showed it is different. For example, Noordzij et al. (18) showed that the expression of CD44v6 mRNA was present in 44% and 75% of the tumors and BPH samples, respectively. But Ni et al. (7) study showed that CD44v6 proteins are highly

expressed in this tissue. In silico analysis was performed the understanding the effect of rs8193C>T in populations on the altered transcription factors or the creation of new locations for transcripts regulatory factors. As shown rs8193C>T loci of CD44v6's, is located in the 3´UTR. Since, it may affect different cell functions, because 3´UTR of gene transcripts is a part of mRNA that is not translated and outright initiated from the translation termination codon, but it has important roles in the structure, function, stability, and expression of the transcript. Also, 3´UTR is mainly a site for the binding of microRNAs and ncRNAs, which regulated gene expression (34). The previous studies showed that the C allele of rs8193C>T might inflict its effect via destabilizing CD44 and miR-570 interaction (25). Whereas in silico studies showed that rs8193 C to T, gain four miRNAs and loss six miRNAs, probably. Also, results showed that rs8193C>T in CD44v6 transcript changed the RNA secondary structure and altered the expression level CD44v6 rs8193C>T transcript may be caused by RNA secondary structure or/and gain and loss of miRNAs.

Since rs8193C>T transition leads to the gain of a new regulatory element site, for Pum2 protein, and loss of YBX1, in the CD44s transcripts. Pum2 is a sequence-specific RNA-binding protein but YBX1 is a DNA- and RNA-binding protein. On the other hand, YBX1 binding to granulocyte-macrophage colony-stimulating factor/GM-CSF mRNA (GM-CSF is a cytokine that promotes eosinophil differentiation) and leads to the stabilization of mRNA and extends cell survival (35), and it can collaborate with ELAV like RNA binding protein 1/ELAVL1, which prevents mRNA decay (36). However, Pum2 mediates post-transcriptional repression of mRNAs via translational inhibition and mRNA degradation (37). Probably the attachment of Pum2 led to degradation of CD44s transcripts and stimulated switch isoforms with overexpression of other variants. The expression of CD44 isoforms affects the Nanog, in the two probably pathways. For example, i. In the Nanog-Stat pathway, Nanog translocates in the nucleus and is coupled with Stat3, and this complex arouses transcription of miR-21 resulting in cell growth and survival in cancer cells. ii. In the PKC-Nanog pathway, CD44 isoforms activate protein kinase C (PKC), and phosphorylated Nanog translocates to the nucleus. The nucleus, creates a ternary complex (p-Nanog, DROSHA, and P68) which affects miR-21 biogenesis. iii. In the PKC-Twist pathway, PKC activation cause Twist phosphorylation and translocation in the nucleus. The p-Twist binding to miR-10 and its transcription result in cell "growth and survival" and "migration and invasion" in cancer cells.

The previous studies showed the downregulation of CD44s in cancer diseases, especially in prostate cancer (38) and loss of CD44 expression was significantly associated with prostate tumor progression (7). Moura et al. (39) demonstrated that prostate neoplasm is caused by a change in the expression of CD44 isoforms which happens with a loss of CD44s and overexpression of all the other CD44 variants. Di Stefano et al. (40) showed that CD44v8-v10 overexpression might be a potential biomarker in gastric and prostate cancer, respectively. Thus, this transition affects the expression of other CD44 transcripts, under the mechanism of isoform switching. By this mechanism expression of CD44 from CD44s switches to CD44v, especially the CD44v8 isoform, because these SNPs are presented in all CD44 variants except for CD44v8.

Based on our findings, the increase of T mutant allele frequency is associated with BPH prostate neoplasm, with OR (0.572), 95% CI (0.364-0.897), and P=0.015. This transition has affected the CD44v6 transcript secondary structure (from -109.30 to -111.30 kcal/mol), miRNA binding sites (gain: four and loss: six for mutant allele), and gain of a new post-transcriptional regulatory element-binding site for Pum2 protein. It seems that the add-on of this site and this protein adjoin to standard CD44 transcript leads to degradation of this transcript, and stimulates switch-isoforms, with the overexpression of other variants. It triggers some of the cell signaling pathways such as Nanog-Stat, PKC-Nanog, and PKC-Twist which could run cell "growth and survival" and "migration and invasion" in cancer cells.

Conclusion

The presence of the T mutant allele in the populations probably can create a new location for regulatory factors on the gene transcripts, which is caused a shift in expression of this protein isoforms. Based on this study, the presence of the T mutant allele of CD44 (rs8193C>T) in the populations may affect transcripts isoform-switching. So, it is known as a prognosis factor and prediction of prostate neoplasms, probably. However, more comprehensive studies in different populations (with various ethnicities and large population sizes), and analysis of *CD44* gene expression in the protein and transcript levels, are required to confirm our data.

Acknowledgments

We would like to extend our thanks to the University of Mazandaran (Iran) for the financial support. There is no conflict of interest to declare.

Authors' Contributions

A.H.C.; Conceptualization, methodology, and software. M.H.; Data curation, writing and original draft preparation. E.M.; Sample preparing and data analysis. All authors were involved in writing this manuscript and approved the final manuscript.

Reference

1. Miah S, Catto J. BPH and prostate cancer risk. Indian J Urol. 2014; 30(2): 214-218.
2. Rawla, P. Epidemiology of prostate cancer. World J Oncol. 2019; 10(2): 63.
3. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature. 1998; 396 (6712): 643-649.
4. Wiklund F, Jonsson BA, Brookes AJ, Strömqvist L, Adolfsson J,

Emanuelsson M, et al. Genetic analysis of the RNASEL gene in hereditary, familial, and sporadic prostate cancer. Clin Cancer Res. 2004; 10(21): 7150-7156.
5. Lai J, Kedda MA, Hinze K, Smith RL, Yaxley J, Spurdle AB, et al. PSA/KLK3 AREI promoter polymorphism alters androgen receptor binding and is associated with prostate cancer susceptibility. Carcinogenesis. 2007; 28(5): 1032-1039.
6. Chen R, Ren S, Sun Y. Genome-wide association studies on prostate cancer: the end or the beginning? Protein Cell. 2013; 4(9): 677-686.
7. Ni J, Cozzi PJ, Hao JL, Beretov J, Chang L, Duan W, et al. CD44 variant 6 is associated with prostate cancer metastasis and chemo-/radioresistance. Prostate. 2014; 74(6): 602-617.
8. Goodison S, Urquidi V, Tarin D. CD44 cell adhesion molecules. J Mol Pathol. 1999; 52(4): 189-96.
9. Gallatin WM, Wayner EA, Hoffman PA, St John T, Butcher EC, Carter WG, et al. Structural homology between lymphocyte receptors for high endothelium and class III extracellular matrix receptor. Proc Natl Acad Sci. 1989; 86(12): 4654-4658.
10. Heldin P, Kolliopoulos C, Lin CY, Heldin, CH. Involvement of hyaluronan and CD44 in cancer and viral infections. Cell Signal. 2020; 65: 109427.
11. Bellerby R, Smith C, Kyme S, Gee J, Günthert U, Green A, et al. Overexpression of specific CD44 isoforms is associated with aggressive cell features in acquired endocrine resistance. Front Oncol. 2016; 6: 145.
12. Knutson JR, Iida J, Fields GB, McCarthy JB. CD44/chondroitin sulfate proteoglycan and alpha 2 beta 1 integrin mediate human melanoma cell migration on type IV collagen and invasion of basement membranes. Mol Biol Cell. 1996; 7(3): 383-396.
13. Günthert U, Hofmann M, Rudy W, Reber S, Zöller M, Haußmann I, et al. A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. Cell. 1991; 65(1): 13-24.
14. Wang Q, Teder P, Judd NP, Noble PW, Doerschuk CM. CD44 deficiency leads to enhanced neutrophil migration and lung injury in Escherichia coli pneumonia in mice. Am J Pathol. 2002; 161(6): 2219-2228.
15. Jordan AR, Racine RR, Hennig MJ, Lokeshwar VB. The role of CD44 in disease pathophysiology and targeted treatment. Front Immunol. 2015; 6: 182.
16. Estess P, DeGrendele HC, Pascual V, Siegelman MH. Functional activation of lymphocyte CD44 in peripheral blood is a marker of autoimmune disease activity. J Clin Invest. 1998; 102(6): 1173-1182.
17. Jeyapalan Z, Deng Z, Shatseva T, Fang L, He C, Yang BB. Expression of CD44 3′-untranslated region regulates endogenous microRNA functions in tumorigenesis and angiogenesis. Nucleic Acids Res. 2011; 39(8): 3026-3041.
18. Noordzij MA, Van Steenbrugge GJ, Verkaik NS, Schröder FH, Van der kwast T. The prognostic value of CD44 isoforms in prostate cancer patients treated by radical prostatectomy. Clin Cancer Res. 1997; 3(5): 805-815.
19. Heider KH, Sproll M, Susani S, Patzelt E, Beaumier P, Ostermann E, et al. Characterization of a high-affinity monoclonal antibody specific for CD44v6 as candidate for immunotherapy of squamous cell carcinomas. Cancer Immunol Immunother. 1996; 43(4): 245-253.
20. Heider KH, Kuthan H, Stehle G, Munzer G. CD44v6: a target for antibody-based cancer therapy. Cancer Immunol Immunother. 2004; 53(7): 567-579.
21. Prochazka L, Tesarik R, Turanek J. Regulation of alternative splicing of CD44 in cancer. Cell Signal. 2014; 26(10): 2234-2239.
22. Rutnam ZJ, Yang BB. The non-coding 3′ UTR of CD44 induces metastasis by regulating extracellular matrix functions. J Cell Sci. 2012; 125(8): 2075-2085.
23. Verma A, Kapoor R, Mittal RD. Cluster of differentiation 44 (CD44) gene variants: a putative cancer stem cell marker in risk prediction of bladder cancer in north indian population. Indian J Clin Biochem. 2017; 32(1): 74-83.
24. Liu Y, Qing H, Su X, Wang C, Li Z, Liu S. Association of CD44 gene polymorphism with survival of NSCLC and risk of bone metastasis. Med Sci Monit. 2015; 21: 2694-2700.
25. Mokhtarian R, Tabatabaeian H, Saadatmand P, Azadeh M, Balmeh N, Yakhchali B, et al. CD44 gene rs8193 c allele is significantly enriched in gastric cancer patients. Cell J. 2020; 21(4): 451-458.
26. Sambrook J, Russell DW. Molecular cloning: a laboratory manual. 3th ed. New York: Cold Spring Harbor Laboratory; 2013.
27. Gerger A, Zhang W, Yang D, Bohanes P, Ning Y, Winder T, et al. Common cancer stem cell gene variants predict colon cancer recurrence. Clin Cancer Res. 2011; 17(21): 6934-6943.
28. Stremitzer S, Zhang W, Yang D, Ning Y, Stintzing S, Sunakawa Y, et al. Variations in genes involved in dormancy associated with outcome in patients with resected colorectal liver metastases. Annal Oncol. 2015; 26(8): 1728-1733.
29. Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. Nucleic Acids Res. 2008; 36(Database issue): D820-D824.
30. Wu X, He X, Li S, Xu X, Chen X, Zhu H. Long non-coding RNA ucoo2kmd. 1 regulates CD44-dependent cell growth by competing for miR-211-3p in colorectal cancer. PLoS One. 2016; 11(3): e0151287.
31. Winder T, Ning Y, Yang D, Zhang W, Power DG, Bohanes P, et al. Germline polymorphisms in genes involved in the CD44 signaling pathway are associated with clinical outcome in localized gastric adenocarcinoma. Int J Cancer. 2011; 129(5): 1096-1104.
32. Lin X, You X, Cao X, Pan S. Association of single-nucleotide polymorphisms of CD44 gene with susceptibility to breast cancer in Chinese women. Med Sci Monit. 2018; 24: 3077-3083.
33. Deng Y, Chen ZJ, Lan F, He QT, Chen SY, Du YF, et al. Association of CD44 polymorphisms and susceptibility to HBV related hepatocellular carcinoma in the Chinese population. J Clin Lab Anal. 2019; 33(8): e22977.
34. Skeeles LE, Fleming JL, Mahler KL, Toland AE. The impact of 3′ UTR variants on differential expression of candidate cancer susceptibility genes. PLoS One. 2013; 8(3): e58609.
35. Capowski EE, Esnault S, Bhattacharya S, Malter JS. Y box-binding factor promotes eosinophil survival by stabilizing granulocyte-macrophage colony-stimulating factor mRNA. J Immunol. 2001; 167(10): 5970-5976.
36. Chen X, Li A, Sun BF, Yang Y, Han YN, Yuan X, et al. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol. 2019; 21(8): 978-990.
37. Van Etten J, Schagat TL, Hrit J, Weidmann CA, Brumbaugh J, Coon JJ, et al. Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. J Biol Chem. 2012; 287(43): 36370-36383.
38. Lou W, Krill D, Dhir R, Becich MJ, Dong JT, Frierson HF Jr, et al. Methylation of the CD44 metastasis suppressor gene in human prostate cancer. Cancer Res. 1999; 59(10): 2329-2331.
39. Moura CM, Pontes J Jr, Reis ST, Viana NI, Morais DR, Dip N, et al. Expression profile of standard and variants forms of CD44 related to prostate cancer behavior. Int J Biol Markers. 2015; 30(1): e49-55.
40. Di Stefano C, Grazioli P, Fontanella RA, De Cesaris P, D'Amore A, Regno M, et al. Stem-like and highly invasive prostate cancer cells expressing CD44v8-10 marker originate from CD44-negative cells. Oncotarget. 2018; 9(56): 30905-30918.