DOI 10.18699/SBB-2022-436

Functional study of rs590352 (*ATXN7L3B*) associated with colorectal cancer

Antontseva E.[1]*, Degtyareva A.[1], Aliev T.[1, 2], Merkulova T.[1, 2]

[1] *Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

* *antontseva@bionet.nsc.ru*

Key words: regulatory SNP, *ATXN7L3B*, colorectal cancer, functional analysis

Motivation and Aim: Previously, using a bioinformatics approach based on the analysis of genome-wide data about allele asymmetry of chromatin protein and transcription factors (TF) binding and allele asymmetry of gene expression, we identified a number of potentially regulatory polymorphisms (rSNPs) [1]. Among them was rs590352 G>C, localized in exon 1 of the *ATXN7LB* gene, for which an association with the risk of developing colorectal cancer was established, the G allele is protective [2]. The aim of the work is an experimental study of the functionality of rs590352.

Methods and Algorithms. Electrophoretic mobility shift assay (EMSA) – method of DNA-probes retention in the gel by nuclear extract proteins was used to study the effect of single nucleotide substitution on TF binding. To search for TFs whose binding sites are affected by single nucleotide substitution, the MotifbreakR software package was used. Oligonucleotides containing experimentally confirmed binding sites for these TFs were used as competitors in EMSA experiments.

The effect of nucleotide substitution on the level of gene expression was studied in a luciferase reporter system. Vector constructs with allelic variants of the inserts were created in pGL4.23[luc2/minP] plasmid. The HepG2 cell line was used for transfection. To assess the allele-asymmetric expression of the *ATXN7LB* gene in human blood mononuclear cells, cDNA libraries of 15 healthy donors heterozygous (rs590352 G/C) were created and sequenced. Sequencing was performed on the Illumina MiSeq platform with ≥1000 reads per library.

Results: In EMSA experiments with a nuclear extract from the HepG2 cell line, it was shown that oligonucleotides V1 (G allele) and V2 (C allele), reproducing allelic variants of the rs590352 location region, have differences in the intensity of the bands on radioautograph, i.e., the G>C nucleotide substitution affects the binding of certain proteins to this region. In cross-competition experiments, when unlabeled V1 and V2 oligonucleotides alternately competed with the labeled DNA probe (V1 or V2), it was shown that a 5-fold and 10-fold excess of competitor V1 attenuates the intensity of DNA probe retention bands by some TFs more strongly than competitor V2, i.e. in the case of the reference allele G, the affinity for some TFs is higher than in the case of allele C. To identify this/these TF by competitive analysis, oligonucleotides were used containing binding sites for: ASCL1, ASCL2, ATOH1, TCF3, TFAP4, ZBTB, MYOG, NHLH1, OLIG2 and PTF1A, predicted using the MotifbreakR software package. It was shown that TCF3 and PTF1A oligonucleotides compete with V1 and V2 for protein binding, in upper and lower bands respectively. Thus, the G>C substitution was shown to weaken TCF3 and PTF1A binding sites. It should be noted that, according to the data of the

ENCODE project, this SNP falls within the ChIP-seq peak for TCF12 (TCF3 paralog), which indirectly confirms the correctness of our result.

It was also shown in the reporter system that luciferase gene expression increases upon transfection with a construct containing double and triple inserts with the G allele. It is of great interest to study the effect of rs590352 on *ATXN7LB* gene expression *in vivo*. In this regard, we evaluated the allele-specific expression of *ATXN7LB* in blood mononuclear cells of heterozygous individuals. It was shown that the expression of the *ATXN7LB* gene in the case of the G allele is 1.3 times higher compared to the C allele.

Conclusion: Thus, our results show that the *ATXN7L3B* gene rs590352, which we previously associated with colorectal cancer, is a regulatory SNP. At the same time, an analysis of the literature data showed that ATXN7LB reduces the activity of the SAGA histone acetyltransferase complex, which implements histone co-transcriptional modifications, leading to an increase in the level of ubiquitinated histone (H2Bub1) and a decrease in transcriptional activity [3]. It has also been shown that H2Bub1 suppresses tumor progression, and a decrease in its level has been observed in the late stages of cancer [4, 5]. Based on our results (improved protein-binding capacity of the G allele, increased reporter gene and *ATXN7LB* gene expression in mononoclears also in the case of the G allele), and given that the G allele was associated with a reduced risk of colorectal cancer, and taking into account that according to the literature increased expression of *ATXN7L3B* is indirectly associated with the suppression of carcinogenesis, the oncoprotective effect of the G allele rs590352 of the *ATXN7L3B* gene can be assumed.

Acknowledgements: This work was supported by the State Budget Project FWNR-2022-0016.

References

1. Korbolina E.E. et al. Novel approach to functional SNPs discovery from genome-wide data reveals promising variants for colon cancer risk. *Hum Mutat*. 2018;39(6):851-85.
2. Leberfarb E.Yu. et al. Potential regulatory SNPs in the ATXN7L3B and KRT15 genes are associated with gender-specific colorectal cancer risk. *Per Med*. 2020;17(1):43-54.
3. Li W. at al. Cytoplasmic ATXN7L3B Interferes with Nuclear Functions of the SAGA Deubiquitinase Module. *Mol Cell Biol*. 2016;36(22):2855-2866.
4. Hahn M.A. et al. The tumor suppressor CDC73 interacts with the ring finger proteins RNF20 and RNF40 and is required for the maintenance of histone 2B monoubiquitination. *Hum Mol Genet*. 2012;21(3):559-68.
5. Urasaki Y. et al. Coupling of glucose deprivation with impaired histone H2B monoubiquitination in tumors. *PLoS One*. 2012;7(5):e36775.