DOI 10.18699/BGRS/SB-2020-255

Associations of methylation level of promoter region of *MLH1* gene and genotypes of the exonic rs1799977

Nadezhda Babushkina
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
nad.babushkina@medgenetics.ru

Anton Markov
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
anton.markov@medgenetics.ru

Irina Goncharova
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
irina.goncharova@medgenetics.ru

Ramil Salakhov
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
ramil.salakhov@medgenetics.ru

Iuliia Koroleva
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
yuliya.koroleva@medgenetics.ru

Anna Postrigan
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
postrigan.anna@medgenetics.ru

Aleksei Zarubin
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
aleksei.zarubin@medgenetics.ru

Aleksei Sleptcov
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
alexei.sleptcov@medgenetics.ru

Aksana Kucher
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
aksana.kucher@medgenetics.ru

Maria Nazarenko
Research Institute of Medical Genetics, TNRMC RAS, Tomsk, Russia
maria.nazarenko@medgenetics.ru

***Abstract* — The eQTL SNP in exon 8 of the *MLH1* gene – rs1799977, may affect the methylation level of the promoter region of the same gene. When studying both the effects of methylation level and individual SNPs on the formation of phenotypic variability in humans and the risk of developing pathological conditions, it is necessary to use a integrated data analysis approach.**

Keywords — MLH1, SNP, CpG-sites, methylation level

Motivation and Aim

DNA methylation is one of the most studied epigenetic mechanisms regulating the transcriptional activity of chromatin. Through methylation/demethylation of DNA, a number of fundamental biological processes are regulated, so the study of methylation profiles in different tissues, in different pathologies, in dynamics in ontogenesis, etc., is currently actively carried out. DNA methylation levels are known to be sensitive to a wide range of external influences (diet, drug intake, level of physical activity, and various environmental factors). But in addition, DNA methylation may also depend on structural variations of the genome. Several studies have shown that the level of DNA methylation depends, for example, on cis-localized regulatory SNP [1].

In this work, a pilot study of the methylation level of the studied regulatory regions of the *MLH1* gene versus genotypes was conducted for one of the SNP eQTL in the gene. *MLH1* encodes one of the miss-match repair system proteins, involved in many biological processes. Methylation of DNA in the region of the *MLH1* gene, has variable character and significantly changes in response to various substances. The significance of methylation level changes is proved in oncogenesis [2]. A direct relationship between an increase in the level of methylation of the *MLH1* gene and a decrease in expression at the mRNA level has been proven [3].

Studing of the dependence of the he methylation level of the promoter region *MLH1* gene with genotypes on the regulatory exonic SNP of the same gene.

Methods

Methylation was studied in leukocytes of peripheral blood of 22 patients with atherosclerosis of carotid arteries (stenosis level >70%) and 14 healthy individuals (stenosis level <28%). The groups examined are comparable by sex and age, represented by Russian residents of Tomsk. Informed agreement on the purpose and possible risks of the study was signed by all participants.

The *MLH1* gene promoter was studied, which is the “coast” of the CpG-island (within ± 2000 bp from the CpG-island), capturing also a fragment of exon 1 of the *EPM2AIP1* gene (GRCh37/hg19; chr 3: 37033249-37033762).

The genotyping rs1799977 was performed by SNaPshot analysis using Multiplex Kit and Primer Focus kit SNaPshot® ("Applied Biosystems," USA), on the ABI Genetic Analyzer 3730 platform. The analysis of the CpG-sites of the MLH1 gene promoter was performed using the bisulfite method sequencing using the technology of high-performance mass parallel sequencing (massive parallel sequencing, MPS) on the MiSeq platform (Illumina, USA); Bisulfite DNA treatment was performed using the EZ DNA Methylation Kit ("Zymo Research," USA); DNA library preparation for MPS - Nextera DNA Flex Library Prep Kit ("Illuminina," USA) [4]. The study was conducted on the basis of the Center for the collective use of scientific research equipment and experimental biological material “Medical genomics” at the Research Institute of Medical Genetics, Tomsk Scientific and Research Center, RAS.

Bioinformatic processing of the results was carried out using the BWA-Meth programs (for mapping readings) [https://github.com/brentp/bwa-meth], MethylDackel (for extracting information on methylation of CpG-sites), as well as in software methylKit package in the environment R. The

coverage level of CpG-sites with read sequences ranged from 300 to 1100 (with an average reading level of 900). The methylation level of the CpG-site was determined as the percentage of readings corresponding to methylated cytosine residues relative to all readings of cytosine at a given genome position.

Non-parametric tests Kruskall-Wallis, Wilcoxon and Mann–Whitney were used for differentially methylated CpG-sites. Statistically significant differences with achieved level of statistical significance $p<0.05$ were considered, of false-positive control results was carried out by the method of Benjamin-Hochberg.

Results

In the regulatory region of the *MLH1* gene, 9 CpG-sites were analyzed (site coordinates on chromosome 3 (according to assembly GRCh37/hg19): chr3: 37033323-37033662). The "pattern" of methylation of CpG-sites in the studied region of the MLH1 promoter was similar in all samples: the 2nd (chr3: 37033373) and 5th (chr3: 37033562) CpG-sites with a methylation level reduced relative to other CpG-sites flanked areas with higher methylation. This "pattern" approximately repeated the shape of the distribution density of the modified histone H3K27Ac sites in different cell lines.

Significant ($p<0.03$) associations of methylation level with genotypes in rs1799977 were recorded for all analyzed CpG-sites, the differences being more pronounced in the control group. In all cases, carriers of genotype AG had higher levels of DNA methylation (on average about 10%) than carriers of genotype AA. Only one of those examined (from the group of patients with atherosclerosis) was a carrier of the GG genotype, its methylation level of all sites being higher than that of carriers of other genotypes. Single observation cannot make a clear conclusion about the effect of the genotype on the level of DNA methylation. However, this fact is worth noting because it is consistent with the general trend of the level of CpG-site methylation depending on the number of G alleles in the genotype of the surveyed (AA<AG<GG).

When comparing DNA methylation levels between groups of patients and healthy individuals, no statistically significant differences were detected, either in general or taking into account genotypes. At once, for all the studied CpG-sites, AA genotype carriers in the control group had lower methylation levels than patients, while AG genotype carriers in the control group had higher methylation levels than patients. Thus, differences in DNA methylation levels in healthy individuals with different genotypes in rs1799977 significantly exceed those in atherosclerosis patients (5-13% for different sites, $p = 0.00004$).

How the methylation level of the promoter region of a gene (chr3: 37033249-37033762) is associated with genotypes for rs1799977 - a missense substitution in exon 8 (chr3: 37053618), remains unclear. There are two possible explanations for this fact. First, the studied SNP may be in linkage disequilibrium with any SNP in the regulatory region of this gene. Indeed, rs4647200 is resided in the region of the studied CpG-sites (chr3: 37033462). Analysis of bisulfite sequencing data revealed 4 heterozygous and one homozygous carrier of rare allele G by rs4647200 (all in the group of patients with atherosclerosis only). When analyzing bisulfite sequencing data, 4 heterozygous and one homozygous carriers of the rare G allele were identified by rs4647200 (all in the group of patients with atherosclerosis only). There is linkage disequilibrium between the markers rs4647200 and rs1799977 (D'= 1.000, r2 = 0.377). Second, according to the "GeneHancer" database, both the promoter region of the *MLH1* gene and rs1799977 are located in the sites of interaction with the transcriptional factor LRRFIP2 [http://www.genome.ucsc.edu/], which may cause of functional linkage between them. The findings are new information and require further study.

Thus, a pilot study found that rs1799977 in the exon 8 of the *MLH1* gene could have an effect on the methylation level in the promoter region of the same gene. This, in turn, indicates the relevance of using an integrated approach in the analysis both the effects of methylation levels and individual SNPs on the formation of phenotypic variability in humans and the risk of developing of pathological conditions.

Acknowledgment

The work was performed as part of the State Assignment of the Ministry of Science and Higher Education No. 075-00603-19-00.

References

[1] Ahsan M. et al. (2017) The relative contribution of DNA methylation and genetic variants on protein biomarkers for human diseases. PLoS Genet. 13(9): e1007005.

[2] Hitchins M.P. et al. (2011) Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5'UTR. Cancer Cell. 20(2): 200–213.

[3] Wu C.X. et al. (2018) Changes of DNA repair gene methylation in blood of chronic fluorosis patients and rats. J Trace Elem Med Biol. 50: 223-228.

[4] Masser D.R. et al. (2015) Targeted DNA methylation analysis by Next-generation sequencing. J Vis Exp. – 96: e52488/