High intra- and interindividual variations in the expression levels of the human O6-methylguanine-DNA methyltransferase (MGMT) gene have been observed. This DNA repair enzyme can be a cause of resistance of cancer cells to alkylating chemotherapy. It has been studied the association of single nucleotide polymorphisms (SNPs) of MGMT with the risk for different types of cancer, progression-free survival in patients with cancer treated with alkylating chemotherapy, as well as an effect of SNPs on the MGMT gene expression and activity of the enzyme. SNPs have been suggested to be the factors which influence the levels of interindividual variability of the MGMT expression. Therefore, the aim of this paper was to review the experimental data on SNPs of the human MGMT gene, which are associated with cancer, as well as on location of MGMT-SNPs in regulatory and protein-coding regions of the gene in relation to its regulation. Lots of MGMT SNPs, which could affect the gene expression and result in interindividual MGMT variability or the enzyme resistance to pseudosubstrate inhibitors, have been revealed within the promoter and enhancer regions, the 5'- and 3'-UTRs and introns of the MGMT gene, as well as within the protein-coding region. Many of them may have regulatory effect.

Keywords: O6-methylguanine-DNA methyltransferase (MGMT), gene regulation, single nucleotide polymorphism (SNP), transcription factor binding site.

Introduction. The expression of the O6-methylguanine-DNA methyltransferase (MGMT), the DNA repair enzyme responsible for removing alklylation adducts from the O6-guanine in DNA, and its activity determine cell response to alkylating agents, including anticancer chemotherapy, preventing mutations and cell death. This enzyme can provide resistance of cancer cells to such therapy. The high intra- and interindividual variations in the MGMT expression levels have been observed, indicating to a complicated regulation of this gene [1, 2].

It has been suggested that single nucleotide polymorphisms (SNPs) are the factors which influence the levels of interindividual variability of the MGMT expression [3, 4]. The association of some MGMT polymorphisms with the risk for different types of cancer, as well as the effects of polymorphic variations on the gene expression and activity of this enzyme are discussed in [3]. Many known SNPs of the human MGMT gene, which could affect the expression and result in interindividual MGMT variability or the resistance of the enzyme to pseudosubstrate inhibitors, have been revealed within the promoter and enhancer regions, the 5'- and 3'-UTRs and introns of the MGMT gene, as well as within the protein-coding region [3, 4]. It has been shown that at least two intragenic SNPs have the influence upon an interindividual variation of the MGMT activity in peripheral blood mononuclear cells [5], as well as in normal human lung tissue [6].

Currently, disease- and trait-associated SNPs are rapidly being identified in the genome wide association studies and using related strategies [7]. The majority (~93 %) of these variants lie within non-coding sequences, which are concentrated in regulatory DNA marked by deoxyribonuclease I (DNaseI) hypersensitive sites (DHSs) [8]. Thus, SNPs in functionally important non-coding DNA regions could make a significant
contribution to the phenotypic variation and disease susceptibility among individuals [9]. Evaluation of extent and number of SNPs detected within the human MGMT gene can be done by looking at the track of all SNPs in the protein-coding and non-coding regions including promoter from the UCSC genome browser (the University of California Santa Cruz, Figure in Supplement). Since disease-associated SNPs systematically perturb the transcription factor (TF) recognition sequences and frequently alter allelic chromatin states, it is suggested the involvement of some SNPs in transcriptional regulatory mechanisms, including modulation of promoter and enhancer elements and enrichment within the expression quantitative trait loci [8]. SNPs in the regulatory non-coding regions can influence the DNA helical conformations, protein binding, methylation status of CpG dinucleotides, transcript splicing, etc.

The aim of this paper is to review the experimental data on SNPs of the human MGMT gene, which are associated with cancer, as well as on location of MGMT-SNPs in regulatory and protein-coding regions of the gene in relation to its regulation. This article is the second part of a thematic series on the regulation of MGMT expression.

SNPs in regulatory regions of the human MGMT gene. There are several known SNPs in the promoter, enhancer region, and introns of the human MGMT gene. They include G135T, G290A, C485A, C575A, G666A, C777A, G795C, A1034G and C1099T (numbered according to the X61657) [5, 10–12]. The X61657 is a genomic clone 1157 base pairs in length which contains the sequence of the 5'-upstream region of the human MGMT gene with the promoter activity and the first untranslated exon, taken from the NIH genetic sequence database GenBank [13, 14]. We have aligned X61657 to chromosome 10 by using Nucleotide Blast program (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and BLAT alignment tool [15] (Fig. 1).

368

Fig. 1. Single nucleotide polymorphisms (SNPs) from ENCODE [18] within promoter region of the human MGMT gene. The open chromatin region of MGMT promoter, marked by DHS and overlapped with TF ChIP-seq binding sites, is separated by straight lines: 1 – the human MGMT promoter X61657 from BLAT search; 2 – RefSeq gene and TSS location; 3 – DNaseI hypersensitive regions, marked as gray and dark boxes, darkness of which is proportional to the maximum signal strength observed in any cell line (the number to the left of the box shows how many cell lines are hypersensitive in the region); 4 – the track of TF ChIP-seq, which shows regions where TFs bind to DNA as assayed in different cell lines (the darkness of the box is proportional to the maximum signal strength observed in any cell line); 5 – CpG island; 6 – repeating elements (the open chromatin region contains C-rich repeat (Low complexity Family and Class), 156 bps in length at position chr10:131265319-131265474 on plus strand of DNA); 7 – single nucleotide polymorphisms track from dbSNP build 137 which have a minor allele frequency of at least 1 %, mapping only once to reference assembly.
The changes within the promoter and enhancer regions can affect the gene expression potentially, but it is difficult to check this, because many other factors can influence the \textit{MGMT} expression. In particular, it is suggested that the C575A variant revealed in melanoma patients could influence the gene transcription [10], as it is located closely to the binding site (BS) with the eukaryotic heat shock factor [13]. The C1099T polymorphism located within the enhancer region in the exon 1 [16] has been shown to increase the \textit{MGMT} promoter-enhancer activity in luciferase assay [11].

Since recent numerous studies have identified many disease- and trait-related genetic variants [7], we analysed the human \textit{MGMT} gene non-coding and coding regions for SNP enrichment. The Fig. 1 presents a track that contains the information about SNPs located within an open chromatin area marked by DHS and overlapped with TF ChIP-seq BSs [17, 18]. They include rs1625649; rs35322871; rs113813075; rs79442343; rs34180180; rs79442343; rs112837630; rs34138162; rs1623007; rs2782888; rs181536588; rs16906252; rs113327489; rs186050433; rs16906255; rs149452540 (Table 1). The location and data on SNPs were taken from the UCSC Genome Browser, the SNP database (dbSNP, build 137) at NCBI [17] and the SNPedia resource, which is focused on the medical, phenotypic and genealogical associations of SNPs [19].

Some of the revealed SNPs are associated with cancer. In particular, the TT genotype of rs1625649 located 523 bases upstream the transcription start site (TSS) of \textit{MGMT} was found to correlate with a worse progression-free survival in patients with metastatic colorectal cancer treated with oxaliplatin-based chemotherapy than the combined GG + GT genotypes [20]. The coding-synonimous polymorphism R (CGC) \textrightarrow (CGT) in NM_002412: synonymous_variant R (CGC) \textrightarrow R (CGT) (Table 1). The location and data on SNPs were taken from the UCSC Genome Browser, the SNP database (dbSNP, build 137) at NCBI [17] and the SNPedia resource, which is focused on the medical, phenotypic and genealogical associations of SNPs [19].

Table 1

<table>
<thead>
<tr>
<th>SNP ID from dbSNP/137</th>
<th>Position</th>
<th>Strand</th>
<th>Observed</th>
<th>Reference allele</th>
<th>Allele frequencies</th>
<th>UCSC’s predicted function relative to MGMT gene tracks</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1625649</td>
<td>chr10:131264931</td>
<td>–</td>
<td>G/T</td>
<td>T</td>
<td>G: 61.53 %; T: 38.47 %</td>
<td>upstream_gene_variant (523 bases)</td>
</tr>
<tr>
<td>rs35322871</td>
<td>chr10:131264957</td>
<td>+</td>
<td>/G</td>
<td>G</td>
<td>Not indicated</td>
<td></td>
</tr>
<tr>
<td>rs113813075</td>
<td>chr10:131265021</td>
<td>+</td>
<td>A/C</td>
<td>C</td>
<td>A: 2.78 %; C: 97.22 %</td>
<td></td>
</tr>
<tr>
<td>rs79442343</td>
<td>chr10:131265102</td>
<td>+</td>
<td>A/G</td>
<td>G</td>
<td>A: 4.39 %; G: 95.61 %</td>
<td></td>
</tr>
<tr>
<td>rs34180180</td>
<td>chr10:131265112</td>
<td>+</td>
<td>A/G</td>
<td>G</td>
<td>A: 3.42 %; G: 96.58 %</td>
<td></td>
</tr>
<tr>
<td>rs189357135</td>
<td>chr10:131265135</td>
<td>+</td>
<td>A/G</td>
<td>G</td>
<td>A: 0.41 %; G: 99.59 %</td>
<td></td>
</tr>
<tr>
<td>rs112837630</td>
<td>chr10:131265201</td>
<td>+</td>
<td>G/T</td>
<td>G</td>
<td>Not indicated</td>
<td></td>
</tr>
<tr>
<td>rs34138162</td>
<td>chr10:131265223</td>
<td>+</td>
<td>A/C</td>
<td>C</td>
<td>A: 5.35 %; C: 94.65 %</td>
<td></td>
</tr>
<tr>
<td>rs1623007</td>
<td>chr10:131265287</td>
<td>–</td>
<td>A/T</td>
<td>T</td>
<td>A: 3.11 %; T: 96.89 %</td>
<td></td>
</tr>
<tr>
<td>rs2782888</td>
<td>chr10:131265328</td>
<td>–</td>
<td>A/C</td>
<td>A</td>
<td>C: 0.091 %; G: 96.98 %; T: 2.93 %</td>
<td></td>
</tr>
<tr>
<td>rs181536588</td>
<td>chr10:131265388</td>
<td>+</td>
<td>A/T</td>
<td>T</td>
<td>A: 10.03 %; T: 89.97 %</td>
<td></td>
</tr>
<tr>
<td>rs16906252</td>
<td>chr10:131265545</td>
<td>–</td>
<td>A/G</td>
<td>G</td>
<td>A: 3.42 %; G: 96.58 %</td>
<td></td>
</tr>
<tr>
<td>rs113327489</td>
<td>chr10:131265576</td>
<td>+</td>
<td>C/G</td>
<td>G</td>
<td>Not indicated</td>
<td></td>
</tr>
<tr>
<td>rs186050433</td>
<td>chr10:131265642</td>
<td>+</td>
<td>C/T</td>
<td>C</td>
<td>C: 99.13 %; T: 0.87 %</td>
<td></td>
</tr>
<tr>
<td>rs16906255</td>
<td>chr10:131265665</td>
<td>–</td>
<td>A/C</td>
<td>A</td>
<td>A: 8.02 %; C: 91.98 %</td>
<td></td>
</tr>
<tr>
<td>rs149452540</td>
<td>chr10:131265680</td>
<td>+</td>
<td>A/C</td>
<td>C</td>
<td>A: 0.96 %; C: 99.04 %</td>
<td></td>
</tr>
</tbody>
</table>
a strong association with the promoter methylation (silencing) in colorectal cancer [21, 22]. Also, it has been shown, that polymorphism C485A in promoter was associated with increased risk of lung cancer in a Korean population, despite this SNP did not have an effect on the promoter activity [12].

Many other SNPs have been also found in introns (Suppl. Fig.). It has been shown that the A allele of the rs7087131 variant of the MGMT gene (chr10:131474474, plus DNA strand; ref. allele G) was associated with a decreased risk of esophageal squamous cell carcinoma [23], while rs12268840 (chr10:131325299, in intron 1, plus DNA strand) – with increased risks of adenocarcinoma of the esophagus [24].

We discussed in our previous paper [25] that promoter of the human MGMT gene from GenBank (X61657) contained TSS, which is located within the CpG island (CGI), DHS and exon 1 of the gene (Fig. 1). The overlapping of CGI and DHS marks the open chromatin region and the location of active cis-regulatory elements. The ENCODE studies of different cell lines have demonstrated by using ChIP-seq (chromatin immunoprecipitation with antibodies specific to the TF followed by sequencing of the precipitated DNA), that the MGMT promoter can be targeted by several TFs in open chromatin region including c-Myc, E2F6, Egr-1, ELF1, HEY1, Nrf1, Oct-2, Pol2, POU2F2, TAF1 (Fig. 1, Table 2). TFBSs form two clusters located close to each other. It has been identified a 59 bp enhancer within the second cluster, which produced increased transcriptional activity in a reporter gene assay and was required for efficient promoter function [16]. This enhancer sequence is located at the first exon/intron at position +144 to +202 with respect to the TSS of the MGMT promoter X61657 [16]. This sequence overlaps with BSs for c-Myc, Oct-2, Pol2, POU2F2, TAF1 revealed in different cell types (Fig. 1).

Since SNPs can perturb TF recognition sites, we have analyzed the DNA sequences around polymorphic nucleotides ± 20 bases for any changes (Fig. 2) using Match, which utilizes a library of positional weight matrices from TRANSFAC® Public 6.0 (http://www.gene-regulation.com/pub/programs.html#match), as well TFSEARCH program (http://www.cbrc.jp/research/db/TFSEARCH.html). It has been revealed using Match, that polymorphic A allele of rs34180180 caused disappearance of BS for C/EBP (CCAAT/enhancer binding factor), and polymorphic C allele of rs16906255 for-

\[ \text{A} \quad \text{Reference allele} \]  
\[ \text{Polymorphic allele} \]

<table>
<thead>
<tr>
<th>rs34180180</th>
<th>GAAAGGTAC GGGCCATTGT C CAAACTAGG CACAGGCTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>V6CEP3_01</td>
<td>14 (-) 0.977 0.975 cagtttgCACA Ac C/EBP*</td>
</tr>
<tr>
<td>V6CEP3_02</td>
<td>14 (-) 0.984 0.986 cagtttgCACA Ac C/EBP</td>
</tr>
<tr>
<td>rs16906255</td>
<td>GTGGGGGCTCA GGGCCCGCCC # CACCCCGAGG GGGGAGG</td>
</tr>
<tr>
<td>V6SAP2_04</td>
<td>16 (+) 1.000 0.986 gggcccCACCAG Gacx-4</td>
</tr>
<tr>
<td>V6SAP2_06</td>
<td>22 (+) 0.942 0.947 cAGCCCGAaggg Aap-2</td>
</tr>
<tr>
<td>rs189357135</td>
<td>AACTAGGCA CACAGGCTT  G GGGGAAGG CCGAGGCC</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00148 SKY</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00877 Ik-2**</td>
</tr>
<tr>
<td>rs2782888</td>
<td>GAGCTGGGGC GGGCCCGGG # GGGGGCCGAT GTAACCACCC</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00088 Sp1</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00853 c-Rel</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00877 Ik-2</td>
</tr>
<tr>
<td>rs186050433</td>
<td>GCCTCAAGG GTCTGGAGG G GTCTCACTT CCGCTGAGG</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00271 AML-1a</td>
</tr>
</tbody>
</table>

\[ \text{B} \quad \text{Polymorphic allele} \]

<table>
<thead>
<tr>
<th>rs34180180</th>
<th>GAAAGGTAC GGGCCATTGT C CAAACTAGG CACAGGCTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>V6CEP3_01</td>
<td>14 (-) 0.977 0.975 cagtttgCACA Ac C/EBP*</td>
</tr>
<tr>
<td>V6CEP3_02</td>
<td>14 (-) 0.984 0.986 cagtttgCACA Ac C/EBP</td>
</tr>
<tr>
<td>rs16906255</td>
<td>GTGGGGGCTCA GGGCCCGCCC # CACCCCGAGG GGGGAGG</td>
</tr>
<tr>
<td>V6SAP2_04</td>
<td>16 (+) 1.000 0.986 gggcccCACCAG Gacx-4</td>
</tr>
<tr>
<td>V6SAP2_06</td>
<td>22 (+) 0.942 0.947 cAGCCCGAaggg Aap-2</td>
</tr>
<tr>
<td>rs189357135</td>
<td>AACTAGGCA CACAGGCTT  G GGGGAAGG CCGAGGCC</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00148 SKY</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00877 Ik-2**</td>
</tr>
<tr>
<td>rs2782888</td>
<td>GAGCTGGGGC GGGCCCGGG # GGGGGCCGAT GTAACCACCC</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00088 Sp1</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00853 c-Rel</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00877 Ik-2</td>
</tr>
<tr>
<td>rs186050433</td>
<td>GCCTCAAGG GTCTGGAGG G GTCTCACTT CCGCTGAGG</td>
</tr>
<tr>
<td></td>
<td>-------&gt; M00271 AML-1a</td>
</tr>
</tbody>
</table>

Fig. 2. Examples of effect of polymorphic allele upon change of recognition motif for TF. Bioinformatic analysis has been done by using Match (A) and TFSEARCH (B) programs. Notes: *matrix identifier, position (strand), core match, matrix match, sequence (always the plus strand is shown), factor name; **TFMATRIX entries with High-scoring
med BS for Pax-4 (member of the paired box family of TFs). Polymorphic A allele of rs189357135 has been shown by using TFSEARCH to create two new BSs for c-Ets-1 and Elk-1 (TCF) TFs, C allele of rs 2782888 – for SP1, and T allele of rs186050433 – for E47 TF. Thus, this short prediction analysis has shown a possible implication of SNPs in the change of recognition sequences for TFs within the human MGMT promoter. Also, SNPs can change the methylation pattern of CGI of promoter and gene body, in particular due to the forming new or destruction of existing CpG dinucleotides in DNA.

For example, in an open chromatin region of the given gene there are three SNPs (rs79442343, rs113327489, rs186050433), a reference allele of which forms CpG dinucleotide (Table 1). The rs2782888 SNP forms such dinucleotide, increasing a total number of CpG dinucleotides within CGI, methylation of which can potentially be attractive to binding with methyl-CpG binding proteins which in turn recruit histone deacetylase complexes and cause the chromatin condensation [26].

Also, we have classified SNPs revealed from the open chromatin area of the human MGMT promoter by using Regulome database [27] to identify a putative regulatory potential and a functional role of variants (Table 3). According to this database, only three of the analyzed SNPs (rs112837630, rs2782888, rs16906255) belong to the group of variants, which likely affect TF binding, since these SNPs are located within TF binding motif and DNase hypersensitive site [27]. However, the Fig. 1 shows higher amount of such SNPs, which are located within an open chromatin area and may affect TF binding. This difference is probably caused by the fact that RegulomeDB contains data from the old built 132 of dbSNP.

SNPs in the protein-coding region of the human MGMT gene. SNPs in the protein-coding region of the MGMT gene were also found out in populations of healthy donors and patients with different types of cancers. For instance, SNP at codon L53L (silent coding effect) in exon 3 was detected in melanoma patients and healthy Swedish individuals [10, 28] and those from the UK [5], L84F (missense) in exon 3 – in melanoma patients and healthy individuals in Sweden [10, 28], lung patients and healthy donors from Poland [11] and the UK [5], I143V (rs2308321, missense) in exon5 – in Swedish [10, 28], Polish [11], English [5] and Caucasian [29] populations. SNP at codon G160R (missense) in exon 5 was detected in young patients with adult type cancers and in the control group [30]. K178R (rs2308327, missense) in exon 5 – in melanoma patients and healthy individuals in Sweden [10, 28], lung patients and healthy controls from Poland [11] and in Caucasian population [29], as well as from the UK [5], I143V (rs2308321, missense) in exon 3 – in melanoma patients and healthy individuals in Sweden [10, 28], lung patients and healthy donors from Poland [11] and the UK [5]. A silent SNP A197A in exon 5 was found in melanoma patients and healthy Swedish individuals [10].

Frequent in the human population polymorphic variants L84F, I143V and K178R are shown to have the same DNA repair efficiency as the wild-type MGMT enzyme, but they are more resistant to the pseudosub-

<table>
<thead>
<tr>
<th>Factor</th>
<th>Position</th>
<th>Genomic Size</th>
<th>Cluster Score (out of 1000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pol2</td>
<td>chr10:131264940-131265389</td>
<td>450</td>
<td>1000</td>
</tr>
<tr>
<td>Pol2 (phosphoS2)</td>
<td>chr10:131265030-131265313</td>
<td>284</td>
<td>351</td>
</tr>
<tr>
<td>Pol2-4H8</td>
<td>chr10:131265045-131265320</td>
<td>276</td>
<td>487</td>
</tr>
<tr>
<td>Pol2</td>
<td>chr10:131265442-131265759</td>
<td>318</td>
<td>918</td>
</tr>
<tr>
<td>Pol2-4H8</td>
<td>chr10:131265444-131265734</td>
<td>291</td>
<td>395</td>
</tr>
<tr>
<td>ELF1_(SC-631)</td>
<td>chr10:131265006-131265281</td>
<td>276</td>
<td>157</td>
</tr>
<tr>
<td>ELF1_(SC-631)</td>
<td>chr10:131265483-131265729</td>
<td>247</td>
<td>367</td>
</tr>
<tr>
<td>Oct-2</td>
<td>chr10:131265419-131265686</td>
<td>268</td>
<td>294</td>
</tr>
<tr>
<td>HEY1</td>
<td>chr10:131265510-131265642</td>
<td>133</td>
<td>242</td>
</tr>
<tr>
<td>HEY1</td>
<td>chr10:131265128-131265249</td>
<td>122</td>
<td>243</td>
</tr>
<tr>
<td>POU2F2</td>
<td>chr10:131265419-131265686</td>
<td>268</td>
<td>293</td>
</tr>
<tr>
<td>c-Myc</td>
<td>chr10:131265081-131265290</td>
<td>210</td>
<td>236</td>
</tr>
<tr>
<td>c-Myc</td>
<td>chr10:131265434-131265643</td>
<td>210</td>
<td>334</td>
</tr>
<tr>
<td>TAF1</td>
<td>chr10:131265086-131265349</td>
<td>264</td>
<td>311</td>
</tr>
<tr>
<td>TAF1</td>
<td>chr10:131265429-131265771</td>
<td>343</td>
<td>259</td>
</tr>
<tr>
<td>Nrf1</td>
<td>chr10:131265139-131265388</td>
<td>250</td>
<td>291</td>
</tr>
<tr>
<td>E2F6_(H-50)</td>
<td>chr10:131265205-131265404</td>
<td>200</td>
<td>84</td>
</tr>
<tr>
<td>Egr-1</td>
<td>chr10:131265216-131265431</td>
<td>216</td>
<td>376</td>
</tr>
</tbody>
</table>
strate inhibitors, such as O\(^6\)-benzylguanine and O\(^6\)-(4-bromothenyl) guanine, or PaTrin-2 [3, 4]. In particular, it has been revealed that I143V and I143V/K178R variants have no effect upon the DNA repair activity compared with the wild-type MGMT [28]. Despite I143V is located closely to the C145 residue at the active site of MGMT, this variant has been shown to have no effect on the activity of the enzyme, but it was more resistant to inactivation by the PaTrin-2 [5]. A higher sensitivity of proteins to inactivation by PaTrin-2 has been also observed for variants I143-K178 and I143-R178 comparable with the V143-K178 and V143-R178 alleles [5].

SNPs L84F and I143V showed a statistically significantly increased risk of lung cancer in Caucasian population, especially in smoking women with non-small cell lung cancer [29], and of adenocarcinoma of the esophagus [24]. In another study, distributions of L53L and L84F polymorphisms did not significantly differ between the lung cancer patients and healthy controls [12]. I143V and K178R have been recently shown to correlate with an increased risk of temozolomide-induced myelosuppression [31]. It has been shown that K178R is associated with increased risks of adenocarcinoma of the esophagus [24], lung cancer risk [32], and with colorectal cancer risk [33]. The association of MGMT polymorphisms with a risk for lung, breast, colorectal and endometrial cancer is reviewed in [3].

**Conclusions.** Many SNPs have been identified within the human MGMT gene coding and non-coding regions, in particular within the promoter region, 5'- and 3'-UTRs and introns. Some SNPs, which are detected in the open chromatin area of promoter region and marked by DHS, overlap with experimentally revealed in different cell lines TF BSs. Consequently, they can perturb these recognition sites, change methylation pattern, due to forming new or destruction of existing CpG dinucleotides in DNA, and as a result, such SNPs may affect the gene expression. The association of MGMT SNPs with the risk for different types of cancer, progression-free survival in patients with cancer treated with alkylating chemotherapy, as well as an effect of SNPs on the gene expression, activity of this enzyme and its resistance to pseudosubstrate inhibitors were discussed. Thus, SNPs can be the factors which influence the levels of interindividual variability of the MGMT expression.

**Acknowledgements.** The study was supported by grant for Young Scientists from the National Academy of Sciences of Ukraine (0111U008220).

---

<table>
<thead>
<tr>
<th>dbSNP ID</th>
<th>Category</th>
<th>Other resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1625649</td>
<td>4*</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs35322871</td>
<td>4</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs113813075</td>
<td>4</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs79442343</td>
<td>4</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs34180180</td>
<td>4</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs112837630</td>
<td>2b**</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs34138162</td>
<td>3a***</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs1623007</td>
<td>4</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs2782888</td>
<td>2b</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs16906252</td>
<td>4</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs113327489</td>
<td>4</td>
<td>UCSC</td>
</tr>
<tr>
<td>rs16906255</td>
<td>2b</td>
<td>UCSC</td>
</tr>
</tbody>
</table>

Category Description: * (4) – Minimal TF binding evidence + DNase peak; ** (2b) – Likely to affect TF binding + any motif + DNase footprint + DNase peak; *** (3a) – Less likely to affect TF binding + any motif + DNase peak.
псевдосубстратных инициаторов включено у промоторов та экзонах, 5' и 3'-нечетных последовательностях и интернах гена MGMT, а также у блок-одноклещей дНК. Благодаря этим факторам возможны следующие регуляторные взаимодействия.

Ключевые слова: О-метилглутамид-DНК метилтрансфераза (MGMT), регуляция экспрессии гена, однонуклеотидные полиморфизма (ОНП), сайт связывания с транскрипционным фактором.

А. П. Яцына, О. В. Пидтала, Л. Л. Лукаш
Экспрессия гена MGMT: понимание ее регуляции.

1. Донкункулоидные полиморфизмы

Резюме

Известно, что существуют значительные внутри- и межиндивидуговые вариации уровней экспрессии гена O-метилглутамид-DНК метилтрансферазы (MGMT) человека. Этот фактор регуляции ДНК может быть причиной устойчивости раковых клеток к алкилирующим химиотерапевтическим. Изучены ассоциации однонуклеотидных полиморфизмов (ОНП) MGMT с риском для различных типов рака, выявляемые пациентов без прогрессирования заболевания после алкилирующей химиотерапии, а также влияние ОНП на экспрессию гена MGMT и активность этого фермента. Высказано предположение, что ОНП являются факторами, влияющими на межиндивидуальные вариации внутриклеточной экспрессии MGMT. В обзоре рассмотрены экспериментальные данные относительно ОНП гена MGMT человека, ассоциированных с раком, а также по локализации ОНП MGMT в регуляторных и белок-кодирующих участках гена, причастных к регуляции. Большое количество ОНП MGMT с потенциальной способностью влиять на экспрессию гена и приводить к межиндивидудуальной вариабельности MGMT или на способность фермента к псевдосубстратным ингибиторам выявлено в промоторной и энхансерной областях, 5'- и 3'-нечетных участках и интронах гена MGMT, а также в белко-кодирующих участках. Многие из них могут служить регуляторными факторами.

Ключевые слова: O-метилглутамид-DНК метилтрансфераза (MGMT), регуляция экспрессии гена, однонуклеотидный полиморфизм (ОНП), сайт связывания с транскрипционным фактором.

REFERENCES


373


