

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Logic of Epigenetics and Investigation of Potential Gene Regions

Metin Budak

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.81739>

Abstract

In living organisms, all molecular structures are formed, and all events are carried out by specific DNA sequences referred to as 'genes'. However, these genes need to be governed and controlled to function properly. In this way, they can participate in biological processes at the optimal time. Genes are actively controlled by other genes and specific proteins called 'transcription factors'. In addition, there is another mechanism that determines gene expression, which can be transmitted from generation to generation and from cell to cell. This mechanism is referred to as the 'epigenetic code'. The DNA sequence does not undergo any changes during the formation of this code, but the relevant part of DNA fragment becomes no longer meaningful. While histone modifications control expression of DNA in chromosome structure, methylation modifications at the gene level are quite effective in controlling expressions the cytosine-end methylation seen in mammalian genome often occurs in the nucleotide pairs which are also called the CpG dinucleotides. The most common epigenetic modifications are the changes in histone proteins and DNA methylation, and the most widely studied and the most well-established epigenetic mechanism is the latter.

Keywords: epigenetic, methylation, mitochondria, methylation-specific PCR, analysis

1. Introduction

Epigenetic changes in living organisms can basically be grouped under two headings. One is protein acetylation, which is an epigenetic modification at the protein level. The other is DNA methylation, an epigenetic modification that occurs at the DNA level. In living species, all macro-molecular structures are determined by nucleotide sequences in the genome, and there is a different mechanism that can be transferred to cell from cell, which has inherent

ability to determine gene expression. It is called epigenetic code. During creation of this code, DNA sequence does not undergo any change. The genetic and epigenetic alterations mentioned above result in the activation of oncogenes or the inactivation of tumour suppressor genes. Methylation may occur in any living organism from bacteria to complex species such as humans. The most common type of methylation is the methylation of gene promoters. This is followed by exon methylation, intron methylation and exon-intron methylations, which may be observed quite frequently.

Methylation-specific PCR (MSP) and methylation-sensitive restriction fragment length polymorphism (MS-RFLP) are the two most widely utilized methods in DNA methylation studies. Also, modified DNA sequencing with bisulphite treatment, known as bisulphite sequencing, may also be employed to investigate the conformation of the region of interest. These 3 are considerably successful methods. With the advances in technology and reduced costs, methylation-specific DNA sequencing has become a frequently used method to investigate the methylated regions identified by means of these methods. Whether a methylation region affects the expression of the gene of interest is another aspect to take into account as some genes may not yield any products although they are not methylated. In that case, one should consider that the gene in question may be activated by other mechanisms. With a better understanding of such histone and DNA modifications, they now attract attention as therapeutic targets in cancer and various diseases. They have started to create new alternatives especially in cancer treatments. Various computer programs have begun to be developed for methylation analysis. This section discusses all of the aforementioned conditions separately.

1.1. Histone modifications

The most basic unit of the structure called chromatin is nucleosomes. A nucleosome is a unit of 146 bp stretch of DNA over H2A, H2B, H3 and H4 central histone proteins and binding of H1 protein to the structure as a lock. In addition, these constructs provide necessary packaging for the DNA molecule, which is quite large, to fit in a small area (**Figure 1**).

Covalent changes in amino acids are found in tail parts of central histone proteins form the epigenetic code. As a result of these changes, chromosome structure constitutes expression control constructs in DNA by acquiring heterochromatin (expressionally inactive) or by forming regions euchromatin (expressionally active regions). Histone modifications can be classified as acetylation, methylation and phosphorylation. The modifications are mostly visible and reversible at the amino (NH_3^-) and carboxyl (COO^-) ends of central histone proteins. Each of these

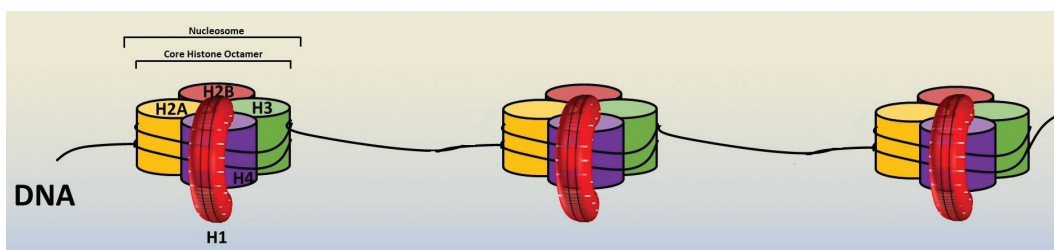


Figure 1. Core histone and nucleosome structure.

changes to square may cause differentiation by altering histone function. For example, it can be seen that one, two and three bases of methyl group are added in methylation of arginine [1]. Histone changes are carried out by various enzymes. These include histone acetyl transferases (HATs), histone deacetylases (HDACs) and histone methyl transferases. The equilibrium in activity of these enzymes and their associated proteins is important when they can perform functions of normal cells. These equilibrium distortions can cause problems that can occur from loss of cell function to cancer formation.

1.2. DNA modifications—methylations

The underlying transcriptional silencing mechanism of DNA methylation is based on the overmethylation of cytosine in CpG-rich islands in the promoter region of a gene. This mechanism cooperates with histone deacetylation to suppress the chromatin structure. GC-rich DNA sequences in the human genome are often found in the promoter region and exon 1 of about 50% of all genes [2]. DNA methylation is the main underlying mechanism that regulates gene expression in mammalian cells, as it happens to be one of the major mechanisms for the silencing of genes involved in cell cycle as well as cell growth and death [3].

The most widely studied and the most well-established epigenetic mechanism is DNA methylation. It is an enzymatic change where cytosines are converted to 5'-methylcytosine. The cytosine-end methylation seen in mammalian genome often occurs at the 5'-CpG-3' dinucleotides, which are also called CpG dinucleotides [4].

Methylation occurs by means of DNA methyltransferase (DNMT) enzymes. The DNMT family consists of four members, namely DNMT1, DNMT2, DNMT3A and DNMT3B. These enzymes are stratified into two groups: those that protect the methylated region and the ones that add new methyl groups. About 70% of all CpG dinucleotides of the human genome are methylated [5]. The remaining are the CpG-rich promoter regions of about 200 base pairs or are the first exons of genes. These regions are also called CpG islands and are found in 60% of all genes [6]. CpG methylation is programmed during the early embryonic period and preserved in later periods. CpG methylation is highly important with regard to normal functions of a given cell, as it affects the regulation of gene expression. For example, DNA methylation plays an important role in gene silencing of the inactive X chromosome as well as the regulation of age-related or tissue-specific gene expression [7].

Although the structural changes that occur in DNA are usually termed as mutations, not every alteration is actually a mutation. A mutation refers to any change at base level such as purine-to-pyrimidine (G-A) or pyrimidine-to-pyrimidine (C-T) changes; single or multiple alterations; insertions, deletions and even single nucleotide polymorphisms (SNP). Yet, SNPs differ from mutations due to their structure. When methylation is compared with other changes in DNA, the methylation process may be considered as another type of mutation, with a change in the structure of the base resulting from a chemical change in DNA. However, mutations are rare changes compared to methylation, and they may or may not be repaired by DNA repair mechanisms [8]. They can be inherited from any ancestor or parent, and they may also occur as germline changes. On the other hand, SNPs can be called DNA alterations, which are more common in the population and which manifest themselves as susceptibility to disease, rather

than resulting in a direct disease phenotype. At this point, methylation is not considered as a mutation, despite the fact that it prevents cytosine behaviour by adding a methyl group from CpG dinucleotides to cytosine [9, 10].

The most appropriate means of this option are the CpG sequences within the DNA, which are bound to their conjugates through an enzymatic process in a stronger manner compared to the A-T pairs. This is because ApTs are bound to their complementary pairs in the corresponding chain by means of two hydrogen bonds, while CpGs are bound with three. Such binding characteristics are expected to provide stability to CpGs compared to ApTs. This may explain the greater frequency of methylation in CpGs rather than ApTs in the organism.

Methylation usually occurs through the addition of a methyl group to CpG sequence or to the C base in these CpG islands. Although such a change normally appears as a mutation, it is understood that, unlike mutations, this change is a highly functional mechanism in terms of cellular development and quite common across living organisms from bacteria to highly complex multicellular species. In this way, the organism can adapt to environmental changes by changing the activation of the desired genes in response to external influences when necessary, thereby maintaining vitality and survival. During a methylation reaction, 5-methylcytosine is formed with the addition of a methyl group to the fifth carbon of the cytosine in CpG base pairs by the DNA methyltransferase enzyme (**Figure 1**). Potentially, any CpG base pair or island may undergo methylation. In addition, the fourth nitrogen of cytosine and sixth nitrogen of adenine, which are usually not found in multicellular organisms, may also be methylated in addition to the 5-methylcytosine formation in bacteria [11].

Genomic imprinting is another example of DNA methylation that is involved in single-allele gene expression. Approximately 80 loci are suppressed in this way. The tissue-specific and condition-specific expressions of these genes occur through the regulation of methylation [12]. At the end of 1970s, a decrease in methylcytosine numbers was observed in the genome of tumour cells [13]. This was referred to as hypomethylation of DNA and was demonstrated in benign and malignant tumours [14]. Hypomethylation of DNA may also activate oncogenes. Studies have shown hypomethylation in *SI00A4*, a metastasis-associated gene in colorectal cancers and the genes, *cyclin-D2* and *maspin*, in gastric carcinomas [14, 15]. Hypomethylation may cause loss of imprinting (LOI), thereby promoting cell proliferation. One of the best examples of this process is the loss of imprinting in the *IGF2/H19* region, which is seen in about 40% of colorectal cancers [1].

2. Bacterial epigenetic mechanisms

Modulation of chromosome organisation is one of the host defence mechanisms against bacterial attacks in eukaryotes. The host cell can often resist bacteria through these highly special and successful defence mechanisms. However, bacteria also have mechanisms that are developed against this system. Some bacteria may contain eukaryote-like proteins and eukaryotic histone translation proteins, which target the chromosomal machinery. In this way, they can activate appropriate enzymes in the host [16].

Several bacteria contain an N4-methylcytosine base whose function has not been fully characterised. There are studies indicating that these N4C modifications affect global gene expression in *Helicobacter pylori*, an example of carcinogenic bacteria [17].

Methylation in bacteria is different from eukaryotes in that it is seen in the fourth carbon of the cytosine as well as methylated adenine (N6-methyladenine) in addition to the fourth carbon of cytosine. DNA methylation occurs in bacteria by methyl binding to cytosine C-5 or N-4 and N6-adenine on the DNA methyltransferase enzyme side. N6-methyladenine is found only in bacteria and in less complex eukaryotes, and not in vertebrates [11]. Interestingly, bacteria also contain a restriction modification system that digests DNA methylase to provide protection against foreign DNA. These consist of the restriction enzyme systems called DcM, which recognises the 5-C cytosine, and Dam, which recognises methylated adenine. Of these, the Dam family is the most well-known protein group. The functional domain of Dam is a DNA MTase with an alpha molecule consisting of a polypeptide of 10 amino acids [18].

Similar to eukaryotes, bacteria also have rRNA methylation. The most important aspect of this methylation is that it creates targets for bacterial infections that can cause infection in humans. While promoter methylation is associated with negative expression, this may not always be the case for exon methylation. Still, sometimes exon methylation shows no effect on gene expression. Investigation on genetic mechanisms affecting cardiomyocyte differentiation includes some studies, which show that intragenic methylations create cellular memory through this mechanism, particularly in pluripotent cells [11].

3. Mitochondrial methylation

In all eukaryotic cells, mitochondrion is the most important organelle for cellular energy and the only organelle containing genomic material apart from the nucleus. Owing to its unique and small genome, this organelle exerts certain proteins and RNAs needed for respiratory reactions and cell growth. Together with the nucleus, it is one of the two genetic systems found in the cell. Mitochondrial DNA (mtDNA) has a circular structure and is located inside mitochondrial matrix, bound to the internal membrane. The mtDNA consists of 16,569 base pairs in a loop form, containing a heavy chain (H) and a light chain. This chain structure contains 2 rRNA molecules, 22 tRNA molecules, and 13 genes necessary for oxidative phosphorylation and electron transport (**Figure 2**). A healthy mitochondrion exerts adequate functions by means of certain proteins that are present in the mechanism of oxidative phosphorylation. This genome is about 16.5 kb in humans, and 13 proteins and rRNAs are synthesised from the mitochondrial genome in mammals [19, 20]. Therefore, the slightest change in mitochondrial genome can potentially affect the life of the cell, and thus the organism [21].

As is the case with mutations, methylation is a mechanism that alters the way the genes work together with the diet, drugs and oxidative stress. Methylation profile of human mtDNA starts from the intrauterine period. With the aid of foetal thyroid hormones, mtDNA copy number and mtDNA methylation are regulated by a thyroid-dependent pathway [23]. In addition, mtDNA is also affected by airway pollutants. The elemental carbon present in benzene and exhaust gas in traffic may influence the number of mtDNA copies by means of ribosomal RNA methylation [24].

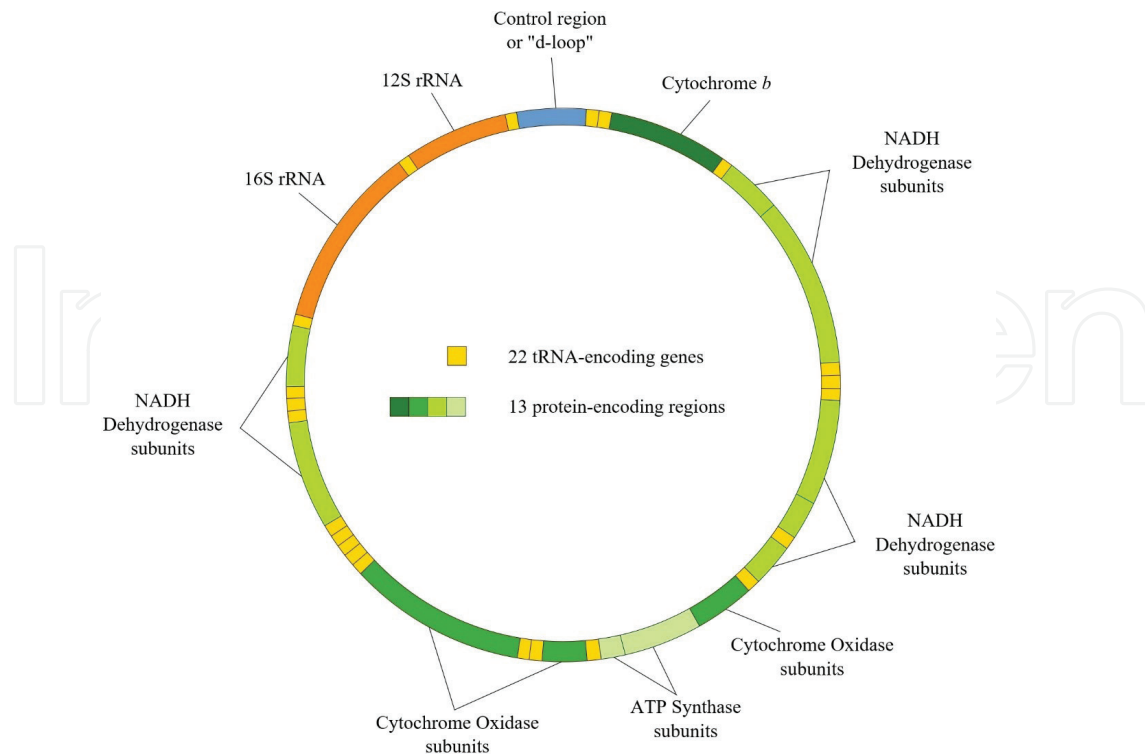


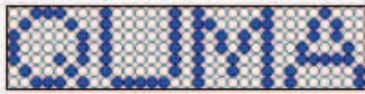
Figure 2. Gene structure of mitochondrial genome [22].

Despite the understanding of these methylation changes in mitochondrial genome, the function of methylated mtDNA has not been fully understood; however, Monique et al. have revealed a different situation. Contrary to what is expected with the methylation of CpGs and GpCs in mtDNA, they have demonstrated that methylation of CpG base pairs had no effect on expression while methylated GpCs were associated with decreased expression [25].

Furthermore, since the mitochondria in humans are entirely of maternal origin, life style of the mother may also have effects at mitochondrion level. Habitual behaviour of the mother, her diet and excessive consumption of fats and sugars trigger obesity, which may affect epigenetics, including that of mitochondria in subsequent generations. A study conducted in mice revealed increased methylation leading to alterations in gene expression and suppression, particularly in the respiratory tract of the offspring of mice that were fed high-fat diets [26]. Furthermore, because the structure of mitochondrion is highly similar to that of bacteria, some genetic factors and structures may also be the same.

4. Detection of the methylation region

Any DNA region containing a CpG sequence may potentially undergo methylation. For this reason, any gene may be subjected to methylation; however, methylation most commonly occurs in the promoter region of genes. That is quite reasonable given the fact that the promoter region is the recognition site for RNA polymerases and therefore of critical importance



QUantification tool for Methylation Analysis
[HOME](#) [OVERVIEW](#) [TERMS OF USE](#) [DOWNLOAD](#) [REFERENCE](#) [LINKS](#)

[日本語](#)
If you have any questions, comments, requests etc., please feel free to contact quma@icb.h.kyushu-u.ac.jp

QUantification tool for Methylation Analysis

You can easily align, visualize and quantify bisulfite sequence data for CpG methylation analysis

☐ [Overview: how to use QUMA](#)
☐ [Quick start](#)
☐ [Execute with sample sequence data](#)
☐ [Sample sequence files](#)

☐ Project name (optional: only contain "A-Z", "a-z", "0-9", "_" and "-") : [?](#)

☐ Enter a target genomic sequence (PCR target) in [plain sequence](#), [FASTA](#) or [GenBank](#) format [?](#)

or upload target genomic sequence (PCR target) file
([plain sequence](#), [FASTA](#) or [GenBank](#) format)

Dosya Seç

Dosya seçilmedi

☐ [Apply sample genomic sequence](#)

☐ Enter bisulfite sequences in [multi-FASTA](#) format [?](#)

or upload bisulfite sequence file
([multi-FASTA](#) format file or [zipped archive of sequence files](#))

Dosya Seç

Dosya seçilmedi

Sequence group name (optional):

☐ [Apply sample bisulfite sequences](#)

☐ **Optional:** Enter second bisulfite sequences for statistical analysis (in [multi-FASTA](#) format) [?](#)

☐ [Apply sample bisulfite sequences](#)

or upload second bisulfite sequence file for statistical analysis
([multi-FASTA](#) format file or [zipped archive of sequence files](#))

Dosya Seç

Dosya seçilmedi

Sequence group name (optional):

☐ [Apply sample bisulfite sequences](#)

☐ Conditions to exclude low quality sequences [?](#)
(CpH: CpA, CpC, CpT)

☐ Conversion [?](#)

☐ Upper limit of unconverted CpHs

☒ C => T conversion
(conversion of forward strand of the genomic sequence)

☐ Lower limit of percent converted CpHs

☐ G => A conversion
(conversion of reverse strand of the genomic sequence)

☐ Upper limit of alignment mismatches

☐ Both
(search both directions and select the best result)

☐ Lower limit of percent identity

☐ Strict CpG site check of bisulfite sequence for repetitive sequence analysis: ☐ [?](#)

☐ Exclude identical bisulfite sequences: ☐ [?](#)

☐ [Hide options](#)

[Go to Home Page](#)

☐ Upload QUMA restore file (.qma) to restore the analysis

Dosya Seç

Dosya seçilmedi

[?](#)

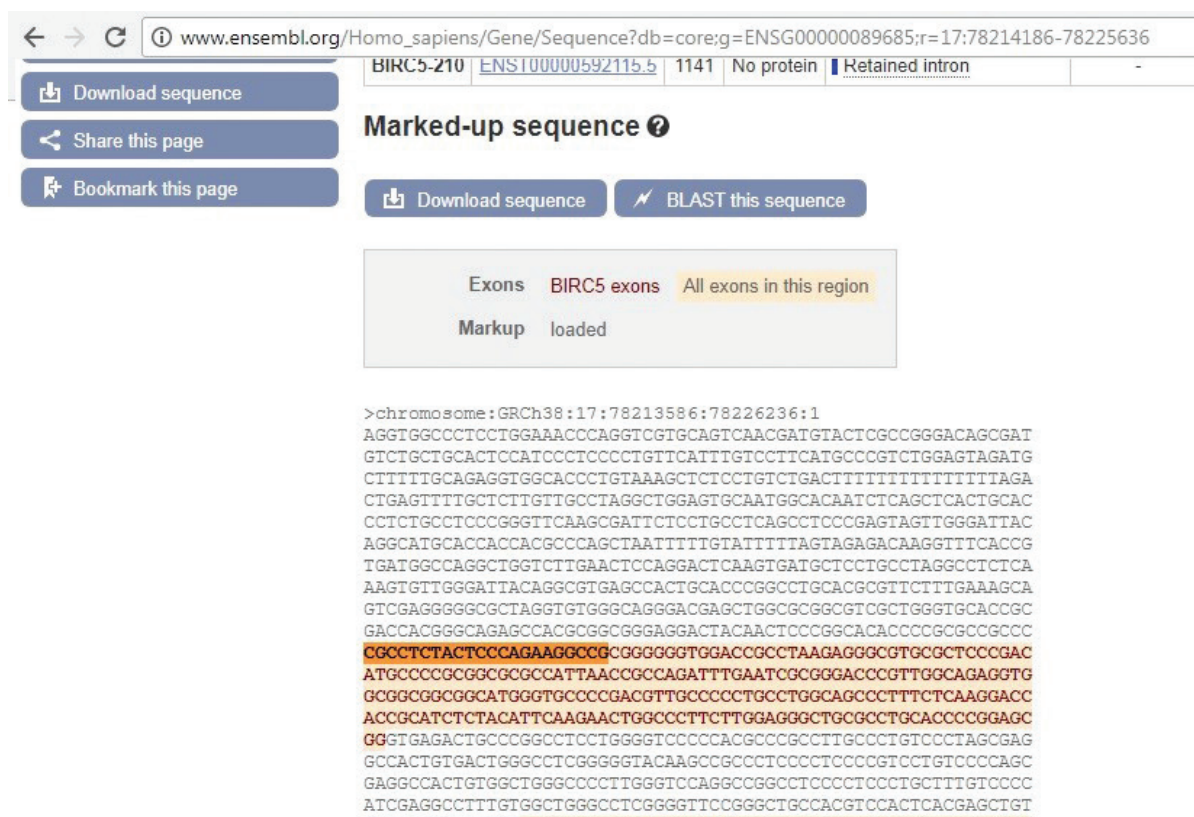
What's New

- Jul 19 2014, Jan 16 2014, Apr 9 2013
 - Minor modification

Figure 3. QUMA Web-based program interface.

for gene expression. Although more rarely, methylation may also be observed in exon 1 and other exons of certain genes. One of the ways of finding out whether a gene or DNA region may undergo methylation is to investigate that region with sequence analysis programs specifically developed for this task. There are paid programs by companies such as Fermentas (Methyl Primer Express® Software v1.0) serving this purpose as well as Web-based free access programs such as 'MethPrimer' developed by Li LC and Dahiya R. (available at <http://www.urogene.org/methprimer/>) [27]. There is also a Web-based program 'DiseaseMeth' (available at <http://202.97.205.78/diseasemeth/Analyze.html#form3>), which offers researchers who is interested in using a study of which disease or cancer is associated with a desired gene methylation. Another Web-based program is SMS (Sequence Manipulation Suite; available at <http://www.bioinformatics.org/sms2/>) [28]. One of many useful programs at this address is program which shows CpG islands in a desired DNA region. Another program is a Web-based program QUMA (Quantification tool for Methylation Analysis; available at <http://quma.cdb.riken.jp/>) (Figure 3) [29]. These programs allow methylation region mapping, designing methylation-specific primers, determining bisulphite sequencing primers, identification of CpG islands and determination of DNA sequences that are altered or newly formed due to bisulphite modifications.

In order to do this, the initial step should be obtaining the FASTA sequence of the sequence in question. The Ensembl genome browser 92, available at <https://www.ensembl.org/>, is a very



The screenshot shows the Ensembl genome browser interface for the Survivin gene (ENSG00000089685) on chromosome 17. The browser displays the gene structure with exons and introns, and the full DNA sequence in FASTA format. The sequence is highlighted in orange, and the gene name 'Survivin' is visible in the header.

Genomic coordinates: 17:78214186-78225636

Gene: BIRC5-210 (ENSG00000089685)

Transcript: ENST00000592115.5 (1141 bp, No protein, Retained intron)

Download sequence | Share this page | Bookmark this page

Marked-up sequence

Download sequence | BLAST this sequence

Exons: BIRC5 exons | All exons in this region

Markup: loaded

```
>chromosome:GRCh38:17:78213586:78226236:1
AGGTGGCCCTCCTGGAAACCCAGGTCGTGCAGTCAACGATGTAAGTCCGCGGGACAGCGAT
GTCTGCTGCACTCCATCCCTCCCTGTTTCAATTTGTCTTCATGCCCCGTCTGGAGTAGATG
CTTTTTCAGAGGTGGCACCCTGTAAAGCTCTCCTGTCTGACTTTTTTTTTTTTTTTTGA
CTGAGTTTGTCTCTTGTTCCTAGGCTGGAGTGCATGGCACAATCTCAGCTCACTGCAC
CCTCTGCCTCCCGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGTTGGGATTAC
AGGCATGCACCCAGCCAGCTAATTTTGTATTTTAGTAGAGACAAGGTTTCACCG
TGATGGCCAGGCTGGTCTTGAAGTCCAGGACTCAAGTGTGCTCCTGCCTAGGCTCTCA
AAGTGTGGGATTACAGGCGTGAGCCACTGCACCCGGCCTGCACGCGTTCTTTGAAAGCA
GTCGAGGGGGCGCTAGGTGTGGGCAGGGACGAGCTGGCGCGGCGTCTGGGTGCACCCG
GACCAAGGGCAGAGCCAGCGCGGGGAGGACTACAACCTCCCGGCACACCCGCGCGCCC
CGCCTCTACTCCAGAGGGCGCGGGGGTGGACCGCCTAAGAGGGCGTGCCTCCCGAC
ATGCCCGCGGGCGGCCATTAAACCGCCAGATTTGAATCGCGGGACCCGTTGGCAGAGGTG
GCGGCGGGCGGATGGGTGCCCCGACGTTGCCCTGCTGGCAGCCCTTTCTCAAGGACC
ACCGCATCTCTACATTCAAGAACTGGCCCTTCTTGGAGGGCTGCACCTGCACCCGGAGC
GGGTGAGACTGCCCGGCTCCTGGGGTCCCCACGCCCGCCTTGCCTGTCCCTAGCGAG
GCCACTGTGACTGGGCTCGGGGTACAAGCCGCCCTCCCTCCCGTCTGTCCCGAGC
GAGGCCACTGTGGCTGGGCCCTTGGGTCCAGGCCGCCCTCCCTCCCTGCTTTGTCCCC
ATCGAGGCTTTGTGGCTGGGCTCGGGGTTCGGGCTGCCACGTCCACTCACGAGCTGT
```

Figure 4. Survivin gene sequence on Ensembl genome browser 92.

The Li Lab
Peking Union Medical College Hospital (PUMCH), Chinese Academy of Medical Sciences

Home Research Publications **Tools & Databases** Protocols People Contact Us

MethPrimer

Invitation to test MethPrimer a.o

Paste an ORIGINAL source sequence. Try this Sample sequence
You don't need to modify your sequence (e.g. convert C to T) before pasting.

```
CCGCTCTACTCCAGAGGCCGCGGGGGTGGACCGCTAAGAGGGCGTGCCTCCCGAC
ATGCCCGCGCGCGCATTAACCGCAGATTTGATCGCGGAGCCGTTGCGAGAGGTG
CGCGCGCGCGCATGGGTGCCCGCGAGCTTGCCTTGCCTGCGAGCCCTTTCTCAAGAGAC
ACCGCATCTCTACATTCAAGAACTGGCCCTCTTGGAGGGCTGCGCTGACCCCGGAGC
GG
```

☒ Pick primers for bisulfite sequencing PCR or restriction PCR.
☐ Pick MSP primers.

☐ Use CpG island prediction for primer selection? Window: 100 Shift: 1 Obs/Exp: 0.6 GC's: 50

Submit Reset

General Parameters for Primer Selection			
Sequence name (optional):			
Target (optional):	"start, size", such as (560, 30)		
Excluded Regions (optional):	"start, size", such as (160, 50 1100, 50)		
Number of output pairs (optional):	5		

Product Size:	Min: 100	Opt: 200	Max: 300
Primer Tm:	Min: 50	Opt: 55	Max: 60
Primer Size:	Min: 20	Opt: 25	Max: 30
Product CpGs:	4	Primer Poly X:	5
Primer non-CpG C's:	1	Primer Poly T:	8

Figure 5. Enter the sequence into Methyl Primer Express. The selection is made for bisulphite sequestration or MSP.

MethPrimer result

Please cite MethPrimer: Li LC and Dahiya R. [MethPrimer: designing primers for methylation PCRs](#). Bioinformatics. 2002 Nov;18(11):1427-31. PMID: [12424112](#)

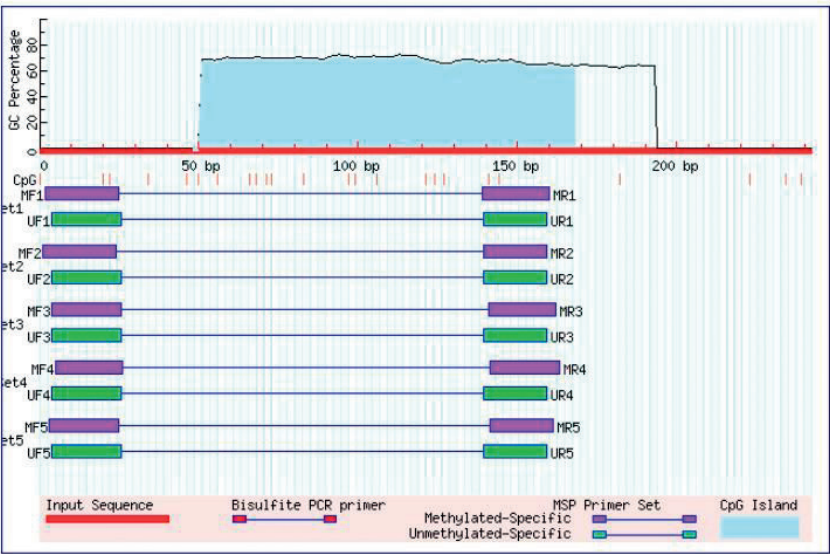


Figure 6. The possible gene region resulting from the analysis.


```

Sequence Name:
Sequence Length: 242

CpG island prediction results
Criteria used: Island size > 100, GC Percent > 50.0, Obs/Exp > 0.60
1 CpG island(s) were found in your sequence
      Size      (Start - End)
Island 1  121 bp    (48 - 168)

Primer picking results for methylation specific PCR (MSP)

  Primer      Start Size  Tm      GC%  'C's  Sequence
1  Left M primer      2   23  58.51  69.57   8  GTTTTATTTTATAGAAAGGTCGCG
   Right M primer    160   21  58.76  76.19   9  CAAACAAAAAACAACGTCGAA
   Product size: 159, Tm: 76.0
   Left U primer      4   22  53.26  68.18   7  TTTTATTTTATAGAAAGGTTGTGG
   Right U primer    159   20  50.93  75.00   9  AAACAAAAAACAACATCAAA
   Product size: 156, Tm: 71.0
2  Left M primer      1   23  58.51  69.57   8  CGTTTTTATTTTATAGAAAGGTCGC
   Right M primer    159   20  55.94  75.00   9  AAACAAAAAACAACGTCGAA
   Product size: 159, Tm: 76.0
   Left U primer      4   22  53.26  68.18   7  TTTTATTTTATAGAAAGGTTGTGG
   Right U primer    159   20  50.93  75.00   9  AAACAAAAAACAACATCAAA
   Product size: 156, Tm: 71.0
3  Left M primer      4   22  59.65  68.18   7  TTTTATTTTATAGAAAGGTCGCGG
   Right M primer    162   21  59.93  76.19   8  ACCAAACAAAAAACAACGTCG
   Product size: 159, Tm: 76.0
   Left U primer      4   22  53.26  68.18   7  TTTTATTTTATAGAAAGGTTGTGG
   Right U primer    159   20  50.93  75.00   9  AAACAAAAAACAACATCAAA
   Product size: 156, Tm: 71.0
4  Left M primer      5   21  58.45  66.67   6  TTTATTTTATAGAAAGGTCGCGG
   Right M primer    163   22  59.94  72.73   8  TACCAAAACAAAAAACAACGTCG
   Product size: 159, Tm: 76.0
   Left U primer      4   22  53.26  68.18   7  TTTTATTTTATAGAAAGGTTGTGG
   Right U primer    159   20  50.93  75.00   9  AAACAAAAAACAACATCAAA
   Product size: 156, Tm: 71.0
5  Left M primer      3   22  57.75  68.18   8  TTTTATTTTATAGAAAGGTCGCG
   Right M primer    161   20  59.10  75.00   7  CCAAACAAAAAACAACGTCG
   Product size: 159, Tm: 76.0
   Left U primer      4   22  53.26  68.18   7  TTTTATTTTATAGAAAGGTTGTGG
   Right U primer    159   20  50.93  75.00   9  AAACAAAAAACAACATCAAA
   Product size: 156, Tm: 71.0

```

Figure 7. Methyl and unmethyl primer sequences and product length for the gene region of interest.

good resource that can be utilised for this step. Ensembl is a database that allows access to DNA sequences in formats such as BLAST and BLAT for comparative genomic studies, evolution studies, sequence variants and transcriptional variants across vertebrate genomes. The relevant DNA sequence obtained from such a database is added to the methylation primer program, the sites of interest are labelled, and the program is run (**Figures 4–8**).

[illegible]

Figure 8. Upper row: original sequence. Lower row: bisulphite modified sequence. (For display, assume all CpG sites are methylated) ++, CpG sites; ::::, non-CpG 'C' converted to 'T'; M>>>>>, left methylated-specific primer; M<<<<<, right methylated-specific primer; U>>>>>, left unmethylated-specific primer; U<<<<<, right unmethylated-specific primer.

5. Methylation-specific PCR (MSP)

MSP is an established and the most commonly utilised method to determine the presence or absence of methylation in a gene region of interest as well as the extent of methylation, if any [30, 31].

In this method, DNA is initially subjected to total bisulphite treatment. In this way, all of the unmethylated cytosines in the DNA sequence are transformed into thymine. However, the methylated cytosines remain unchanged (**Figure 5**). This results in a motif change in the methylated region. Subsequently, spectrophotometric DNA quantification is conducted with DNA samples. For this, measurements are made at wavelengths of 260/280 nm and multiplied by the

dilution factor to determine the amount of DNA as nanogram per microlitre; and according to these amounts, MS-PCR is performed taking care to include equal amounts of DNA samples in the PCR. In this type of PCR, the PCR is conducted in two separate tubes for each sample. While the primer specific for the methylated region is added into one tube, the other tube contains the primer for the unmethylated region, and PCR is performed as 35–40 cycles. PCR samples are then analysed by visualisation with ethidium bromide agarose gel imaging systems.

5.1. Sample reaction: (for the prestin gene promoter region in guinea pigs)

Methylation-specific polymerase chain reaction: DNA purity was measured at wavelengths of 260 and 280 nm on spectrophotometry, and DNA quantification was performed using the DNA ($\mu\text{g/mL}$) = $A_{260} \times \text{Dilution Factor} \times 50$ (coefficient) formula at 260 nm UV. Subsequently, 10 μL of DNA was taken from each sample and bisulphite modification was carried out for DNA with Millipore CpGenome modification kit according to the manufacturer's instructions. This modification converts the cytosines of the unmethylated region to thymines. For the region thought to be altered in this manner, CpG sequences in exon 1 of the prestin gene were detected using the MethPrimer V1.1 beta program [30]. MSP was conducted according to the following protocols to investigate methylation utilising the PCR primers for exon 1 of the prestin gene stated below, and PCR conditions for the methylated and unmethylated regions are as follows:

Methylated region (M: methylated):

M-Forward: ATGTTGAAGAAAATGAAATTTTCGT,

M-Reverse: ACTTATCCCCGATAAAATCCG,

PCR product: 164 base pairs (bp).

Unmethylated region (UnM: unmethylated):

UnM-Forward: TTTATTTTTAGAAGGTTGTGG,

UnM-Reverse: AAACCTACCAAACAAAAACAACATC,

PCR product: 163 bp.

PCR conditions:

PCR buffer 1 \times , MgCl₂: 2.5 mM, DMSO: 5% (v/v), dNTP: 12.5 mM, Primer Forward: 10 pmol, Primer Reverse: 10 pmol, Taq Polymerase: 1 U (5 U/ μL), Template DNA: 100 ng and dH₂O were used to obtain a total of 50 μL , and PCR thermal cycling procedure was as follows:

95°C for 10' (Initiation),

95°C for 45'',	} 35 Cycles
59°C for 30'',	
72°C for 30'',	

72°C for 10'' (Termination).

This investigation allows determining whether the region of interest is methylated and quantifying methylation by measuring the band intensity with any gel analysis system. If desired, results

may also be obtained while real-time PCR is performed with SYBR-green or Taqman or similar fluorescent probes (the probe must be designed according to the bisulphite DNA sequence).

5.2. Evaluation of agarose gel imaging

The resulting primers were stained with ethidium bromide on 2% agarose gel, and agarose gel findings were evaluated by examination under ultraviolet light (Figures 7–10).



Figure 9. Results of polymerase chain reaction in all study groups at the time point of second week for prestin (K1–K3: Group 1; K3–K6: Group 2; T1–T6: Group 3). Complete methylation is seen in sample 6T, while heterozygous methylation is observed in other samples [32].

Example run:



Figure 10. Imaging of MSP and un-MSP PCRs of survivin exon 1. Well 1: marker, 100 bp marker; well 2: positive control; 3–4, 5–6, 7–8, 9–10, 11–12, while unmethylated PCR results are found to be at the same (+) intensity in MSP and un-MSP sample wells of the same cases, the following methylated PCR results were not observed except the control DNA (+), and all samples were accepted as unmethylated. Well 13 is the negative control of the reaction (Figure 10).

5.3. MS RFLP

Another method used to detect any DNA methylation is the methylation-specific restriction fragment length polymorphism (MS-RFLP) method, which produces methylation-specific digestion. This method employs restriction enzymes obtained from bacteria, which recognise

5'	C ↓	C	G	G	3'
3'	G	G	C ↑	C	5'

and cut only specific methylated regions. The most commonly used restriction enzymes for this purpose are HpaII and MspI.

This process digests the methylated CpG (if present) in the DNA region of interest. If that region is not methylated, digestion occurs via HpaII, an isomer of MspI. Concurrent use of these two enzymes provides insight on the methylation status of the region in question.

6. COBRA (combined bisulphite restriction analysis)

COBRA is a method developed by combined use of bisulphite modification and RFLP methods. In this combined method, DNA is first differentiated in the methylation-dependent DNA sequence with sodium bisulphite. As mentioned earlier, in practice, methylase cytosines are not affected, while unmethylated cytosines are converted into uracil. PCR is performed with primers designed specifically for these new DNA sequences obtained by bisulphite method. Unlike MS-PCR, primers used in this PCR step should not contain CpG sequences. After this step, digestion step of restriction comes. At this step, PCR products are treated with two restriction enzymes with TaqI (TCGA) and BstUI (CGCG). These enzymes form a methylation profile by cutting off DNA fragments to properties of whether residues of cytosine are methylated or not (Figure 11).

Methylated regions in DNA fragments digested BstUI and unmethylated homologue of the same region digested TaqI enzyme. These new fragments, which are formed according to methylation state, can be calculated as percentages of methylation rate for region investigated according to patterns and density of band formed by conducting the polyacrylamide gel electrophoresis. It should be noted here that band densities should be determined by a photo



Figure 11. Calculation of band intensity in survivin exon 1 with ‘Gel Quant Express’.

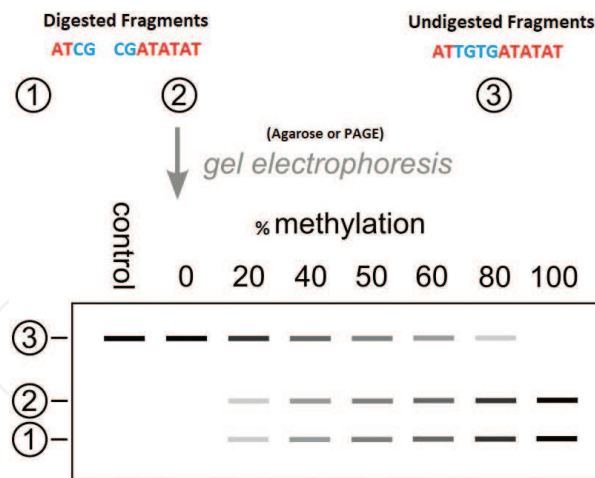


Figure 12. The image of COBRA method and calculation step.

image analysis program (e.g., **Figure 12**). Once these values have been obtained, percentage of methylation can be calculated by substituting the following formula [33, 34] (**Figure 13**).

6.1. Advantages

It is simple, cheap and fast. DNA methylation levels can be shown without needing for any extra bisulphite sequencing.

Due to its high specificity, it is very successful even in DNA material, which is obtained from paraffin blocks.

It is quite quantitative compared to methylation-specific PCR, which is a qualitative method.

In MS-PCR, only locus-specific methylation information can be obtained, whereas in this method, entire region within locus is examined.

6.2. Disadvantages

This method is limited by existing restriction regions in region being investigated.

In addition, incomplete digestions of restriction enzymes can have misleading results for the amount of methylation.

Due to cell-type heterogeneity in different cell complexes, methylated CG sequence may be transformed into other sequences, such as CA or CT, leading to a change in restriction sites.

Considering all these advantages and disadvantages, the COBRA method emerges as an effective method for determining a highly effective level of methylation.

$$\% \text{ Methylation} = 100 \times \frac{\text{Digested Fragments}}{\text{Digested Fragments} + \text{Undigested Fragments}}$$

Figure 13. Mathematical calculation of percentage methylation rates in COBRA method.

7. Conclusion

Many studies after the first discovery of epigenetic changes have shown that epigenetic modifications are quite important in natural flow of life and that many genes are mechanisms used for expression and inactivation when needed.

Furthermore, as these mechanisms are understood better, they have been associated with many pathological conditions, from cancer to mental retardations such as fragile X and Prader-Willi Angelman, to chromosomal instability. In this regard, such diseases have emerged in new therapeutic targets. Chemical agents such as 5-azacytidine and 5-aza-2'-deoxycytidine, which inhibit methylation by binding to and inhibiting DNMT enzyme, are now being tested in phase II-III studies [35–38]. Furthermore, use of oligonucleotides that bind to promoter regions at specific gene level and perform gene inhibition is seen as approaches that may contribute to cancer treatments [39–41].

In addition, histone acetyltransferase inhibitors, which inhibit formation of epigenetic modifications at histone level, have emerged as novel cancer treatment agents. For example, H3-H4 of a soy protein Lunasin has been found to exhibit anticancer properties in mammals by suppressing histone acetylation. Wenyi et al. have shown that YEAST domain, an acetyl lysine-binding module, is effective in the development of cancer, and this domain appears to be the target for anticancer therapies. It seems that such approaches in the future will start to give more successful results [42].

In light of the information presented above, one may conclude that methylation is a highly important genomic mechanism for the cell from unicellular organisms to multicellular organisms. This mechanism is seen in bacteria, mitochondria and all eukaryotic cells in proportion to the complexity and development level of the organism. The identification of methylated regions is as important as the methylation process itself, as this may allow identifying several potential novel targets related to subject matters such as the development mechanism of diseases, certain roles in cancer development and bacterial resistance.

Conflict of interest

There is no conflict of interest.

Author details

Metin Budak

Address all correspondence to: genomicdna2@yahoo.com

Biophysics Department, Faculty of Medicine, Trakya University, Edirne, Turkey

References

- [1] Wong JJJ, Hawkins NJ, Ward RL. Colorectal cancer: A model for epigenetic tumorigenesis. *Gut*. 2007;**56**(1):140-148
- [2] Roman-Gomez J, Jimenez-Velasco A, Agirre X, Prosper F, Heiniger A, Torres A. Lack of CpG island methylator phenotype defines a clinical subtype of T-cell acute lymphoblastic leukemia associated with good prognosis. *Journal of Clinical Oncology*. 2005;**23**(28):7043-7049
- [3] Chim C, Pang R, Liang R. Epigenetic dysregulation of the Wnt signalling pathway in chronic lymphocytic leukaemia. *Journal of Clinical Pathology*. 2008;**61**(11):1214-1219
- [4] Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences*. 2000;**97**(10):5237-5242
- [5] Ehrlich M, Gama-Sosa MA, Huang L-H, Midgett RM, Kuo KC, McCune RA, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Research*. 1982;**10**(8):2709-2721
- [6] Antequera F, Bird A. Number of CpG islands and genes in human and mouse. *Proceedings of the National Academy of Sciences*. 1993;**90**(24):11995-11999
- [7] Plass C, Soloway PD. DNA methylation, imprinting and cancer. *European Journal of Human Genetics*. 2002;**10**(1):6
- [8] Cinemre F, Serinkan B, Cinemre H, Yucel A, Degirmencioglu S, Tuten A, et al. The role of selenoprotein P and selenium in the etiopathogenesis of gestational diabetes mellitus: Association with selenoprotein P1 gene (rs3877899) polymorphism. *Trace Elements and Electrolytes*. 2018;**35**(4):174-182
- [9] Karki R, Pandya D, Elston RC, Ferlini C. Defining “mutation” and “polymorphism” in the era of personal genomics. *BMC Medical Genomics*. 2015;**8**(1):37
- [10] Nei M. The new mutation theory of phenotypic evolution. *Proceedings of the National Academy of Sciences*. 2007;**104**(30):12235-12242
- [11] Sánchez-Romero MA, Cota I, Casadesús J. DNA methylation in bacteria: From the methyl group to the methylome. *Current Opinion in Microbiology*. 2015;**25**:9-16
- [12] Robertson KD. DNA methylation and human disease. *Nature Reviews Genetics*. 2005;**6**(8):597
- [13] Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*. 1983;**301**(5895):89
- [14] Oshimo Y, Nakayama H, Ito R, Kitadai Y, Yoshida K, Chayama K, et al. Promoter methylation of cyclin D2 gene in gastric carcinoma. *International Journal of Oncology*. 2003;**23**(6):1663-1670

- [15] Akiyama Y, Maesawa C, Ogasawara S, Terashima M, Masuda T. Cell-type-specific repression of the maspin gene is disrupted frequently by demethylation at the promoter region in gastric intestinal metaplasia and cancer cells. *The American Journal of Pathology*. 2003;**163**(5):1911-1919
- [16] Rolando M, Gomez-Valero L, Buchrieser C. Bacterial remodelling of the host epigenome: Functional role and evolution of effectors methylating host histones. *Cellular Microbiology*. 2015;**17**(8):1098-1107
- [17] Kumar S, Karmakar BC, Nagarajan D, Mukhopadhyay AK, Morgan RD, Rao DN. N4-cytosine DNA methylation regulates transcription and pathogenesis in *Helicobacter pylori*. *Nucleic Acids Research*. 2018;**46**(7):3429-3445
- [18] Low DA, Weyand NJ, Mahan MJ. Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infection and Immunity*. 2001;**69**(12):7197-7204
- [19] Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. *Nature*. 1981;**290**(5806):457
- [20] Bogenhagen DF. Biochemical isolation of mtDNA nucleoids from animal cells. *Mitochondrial DNA*. 2009:3-14. Springer
- [21] Sun H, Shi W, Wang X. How far can mitochondrial DNA drive the disease? *Mitochondrial DNA and Diseases*. 2017:1-8. Springer
- [22] Falkenberg M, Larsson N-G, Gustafsson CM. DNA replication and transcription in mammalian mitochondria. *Annual Review of Biochemistry*. 2007;**76**:679-699
- [23] Janssen BG, Byun H-M, Roels HA, Gyselaers W, Penders J, Baccarelli AA, et al. Regulating role of fetal thyroid hormones on placental mitochondrial DNA methylation: Epidemiological evidence from the ENVIRONAGE birth cohort study. *Clinical Epigenetics*. 2017;**9**(1):66
- [24] Byun H-M, Panni T, Motta V, Hou L, Nordio F, Apostoli P, et al. Effects of airborne pollutants on mitochondrial DNA methylation. *Particle and Fibre Toxicology*. 2013;**10**(1):18
- [25] van der Wijst MG, van Tilburg AY, Ruiters MH, Rots MG. Experimental mitochondria-targeted DNA methylation identifies GpC methylation, not CpG methylation, as potential regulator of mitochondrial gene expression. *Scientific Reports*. 2017;**7**(1):177
- [26] Keleher MR, Zaidi R, Shah S, Oakley ME, Pavlatos C, El Idrissi S, et al. Maternal high-fat diet associated with altered gene expression, DNA methylation, and obesity risk in mouse offspring. *PLoS One*. 2018;**13**(2):e0192606
- [27] Li L-C, Dahiya R. MethPrimer: Designing primers for methylation PCRs. *Bioinformatics*. 2002;**18**(11):1427-1431
- [28] Stothard P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. 2000
- [29] Kumaki Y, Oda M, Okano M. QUMA: Quantification tool for methylation analysis. *Nucleic Acids Research*. 2008;**36**(suppl_2):W170-W175

- [30] Yalcin O, Budak M. Un-methylation of the survivin gene has no effect on immunohistochemical expression of survivin protein in lung cancer patients with squamous cell carcinoma. *Bratislavské Lekárske Listy*. 2017;**118**(3):160-163
- [31] Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. *Proceedings of the National Academy of Sciences*. 1996;**93**(18):9821-9826
- [32] Bulut E, Budak M, Öztürk L, Türkmen MT, Uzun C, Sipahi T. DNA methylation of the prestin gene and outer hair cell electromotile response of the cochlea in salicylate administration. *Turkish Journal of Medical Sciences*. 2017;**47**(5):1626-1633
- [33] Bilichak A, Kovalchuk I. The combined bisulfite restriction analysis (COBRA) assay for the analysis of locus-specific changes in methylation patterns. *Plant Epigenetics*. 2017:63-71. Springer
- [34] Varinli H, Statham AL, Clark SJ, Molloy PL, Ross JP. COBRA-Seq: Sensitive and quantitative methylome profiling. *Genes*. 2015;**6**(4):1140-1163
- [35] Bauman J, Verschraegen C, Belinsky S, Muller C, Rutledge T, Fekrazad M, et al. A phase I study of 5-azacytidine and erlotinib in advanced solid tumor malignancies. *Cancer Chemotherapy and Pharmacology*. 2012;**69**(2):547-554
- [36] Lin J, Gilbert J, Rudek MA, Zwiebel JA, Gore S, Jiemjit A, et al. A phase I dose-finding study of 5-azacytidine in combination with sodium phenylbutyrate in patients with refractory solid tumors. *Clinical Cancer Research*. 2009;**15**(19):6241-6249
- [37] Levis M, Faderl SH, Garcia-Manero G, Jabbour E, Daver NG, Deltasala SE, et al. Final report of phase II study of sorafenib and 5-azacytidine in patients with relapsed or untreated acute myeloid leukemia and FLT3-ITD mutation. *American Society of Hematology*. 2013;**122**(21):3934
- [38] Kuendgen A, Müller-Thomas C, Lauseker M, Haferlach T, Urbaniak P, Schroeder T, et al. Efficacy of azacitidine is independent of molecular and clinical characteristics-an analysis of 128 patients with myelodysplastic syndromes or acute myeloid leukemia and a review of the literature. *Oncotarget*. 2018;**9**(45):27882
- [39] Urdinguio RG, Sanchez-Mut JV, Esteller M. Epigenetic mechanisms in neurological diseases: Genes, syndromes, and therapies. *The Lancet Neurology*. 2009;**8**(11):1056-1072
- [40] Hua C-D, Bian E-B, Chen E-F, Yang Z-H, Tang F, Wang H-L, et al. Repression of Dok7 expression mediated by DNMT1 promotes glioma cells proliferation. *Biomedicine & Pharmacotherapy*. 2018;**106**:678-685
- [41] Chen M, Nie J, Liu Y, Li X, Zhang Y, Brock MV, et al. Phase Ib/II study of safety and efficacy of low-dose decitabine-primed chemoimmunotherapy in patients with drug-resistant relapsed/refractory alimentary tract cancer. *International Journal of Cancer*. 2018
- [42] de Lumen BO. Lunasin: A cancer-preventive soy peptide. *Nutrition Reviews*. 2005;**63**(1):16-21

