Int J Clin Exp Pathol 2016;9(5):5004-5014
www.ijcep.com /ISSN:1936-2625/IJCEP0023375

*Original Article*

# Genome-wide analysis of differential methylation in pancreatic cancer

Fan Yin[1*], Huan Wang[1*], Yuehua Men[2], Jidong Li[3], Minggen Hu[3], Qingfang Li[1]

*[1]Department of Cancer Center, General Hospital of PLA, 28 Fuxing Road, Beijing 100853, China; [2]Department of Dermatology, Peking University Third Hospital, 49 Huayuan North Road, Beijing 100191, China; [3]Department of Surgical Oncology, Cancer Center, Chinese People's Liberation Army (PLA) General Hospital, 28 Fuxing Road, Beijing 100853, China. [*]Equal contributors.*

Received January 6, 2016; Accepted March 20, 2016; Epub May 1, 2016; Published May 15, 2016

**Abstract:** Pancreatic cancer as a fatal malignant tumor is difficult in diagnosis and treatment. Effective biomarkers are demanded in clinical practice. Up to now, there is little known about the DNA methylation signatures across the whole genome scale in pancreatic cancer. In this study, comparison of differential mathylation sites was performed between pancreatic cancer tissues and pericarcinous tissues using Infinium Human Methylation 450 Beadchips. A total of 24,417 CpG sites representing 9,589 genes were identified between two cohorts. Of the 24,417 CpG sites, 14,721 (60%, 14,721 of 24,417) CpG sites were hypomethylated and 9,705 (40%, 9,705 of 24,417) CpG sites were hypermethylated. GO (Gene Ontology) and KEGG analysis were implemented to systematically characterize the significant differential methylated genes between pancreatic cancer tissues and pericarcinous tissues. In addition, we further screened 51 genes with aberrant methylation, which were the most likely candidate methylation markers within the scale of global differential methylation profiling. GO and KEGG analysis indicated these genes owning a wide range of functions. The identification of differential methylated genes in this study provides information valuable to the in-depth study of pancreatic cancer.

**Keywords:** Pancreatic cancer, DNA methylation, CpG sites, KEGG, GO

## Introduction

The estimated incidence of pancreatic cancer in the United States was 37,700 cases, and an estimated 34,300 patients died from the disease in 2008 [1]. The overall 5-year survival rate among patients with pancreatic cancer is <5% [2]. Pancreatic cancer is more common in elderly persons than in younger persons, and less than 20% of patients present with localized, potentially curable tumors. Several environmental factors have been implicated, but evidence of a causative role exists only for tobacco use. The risk of pancreatic cancer in smokers is 2.5 to 3.6 times that in nonsmokers [3]. Some studies have shown an increased incidence of pancreatic cancer among patients with a history of diabetes or chronic pancreatitis, and there is also evidence that chronic cirrhosis, a high-fat, high-cholesterol diet, and previous cholecystectomy are associated with an increased incidence [4, 5]. More recently, an increased risk has been observed among patients with blood type A, B, or AB as compared with blood type O [6]. In some patients, pancreatic cancer develops as part of a well defined cancer-predisposing syndrome for which germ-line genetic alterations are known. In addition, in some families with an increased risk of pancreatic cancer, a genetic rather than an environmental cause is suspected [7].

Epigenetics is defined as the study of mitotically or meiotically heritable variations in gene function that cannot be explained by changes in DNA sequence [8]. Recently, attention to its role in pancreatic cancer has recently increased. DNA methylation has gained much recent interest for its role in cancer biology. Aberrant patterns of DNA methylation are known to be associated with carcinogenesis and to affect the regulation of genome stability and gene transcription [9]. Genome wide studies of CpG islands have uncovered thousands of loci where differential methylation can segregate pancreatic tumor tissue from normal tissue [10, 11]. Despite this progress, the combination research of changes in DNA methylation in pancreatic

cancer tissues, pericarcinous tissues remains unexplored.

In this study, we have employed a global methylation profiling platform in this work to comprehensively survey a large scale of CpG sites in pancreatic cancer genome. We compared the DNA methylation profiles of the pancreatic tumors and pericarcinous tissues in order to unravel methylation markers for diagnostic purposes. The results suggested that 24,417 CpG sites representing 9,589 genes were detected, among which, 14,721 (60%, 14,721 of 24,417) CpG sites were hypomethylated and 9,705 (40%, 9,705 of 24,417) CpG sites were hypermethylated. We then performed GO (Gene Ontology) and KEGG analysis to systematically characterize the differential methylated genes between pancreatic cancer tissues and pericarcinous tissues. In addition, we further screened 51 genes with aberrant methylation, which were the most likely candidate methylation markers within the scale of global differential methylation. GO and KEGG analysis indicated these genes owning a wide range of functions.

## Materials and methods

### *Subjects and haematoxylin eosin (H&E) staining*

Six patients with pancreatic cancer (2 males and 4 females, mean age: 58.83±14.95 y), without radiation, chemotherapy and immunotherapy treatment, were recruited from General Hospital of PLA in China. The diagnosis of pancreatic cancer was made by at least two experienced oncologists. Sample collection accorded to the following criterions: 1) the minimum diameter of tumor was greater than 2 cm. Meanwhile, pancreatic cancer was identified by haematoxylin and eosin (H&E) staining and the ratio of cancer cells in the whole cells section was over 80%. 2) Tissue adjacent to cancer was collected as far as possible from the cancer tissue in order to avoid the mistake sampling. Pancreatic cancer tissue and tissue adjacent to cancer of each patient were collected and stored in liquid nitrogen immediately for DNA extraction and staining. All specimens were subjected to autolysis for 4 to 8 h and then snap-frozen at -80°C until use in analysis. DNA was extracted from 25 mg samples of the tissue specimens using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's instructions. The DNA yield and purity were determined spectrophotometrically (NanoDrop® ND1000; Thermo Fisher Scientific Inc., Waltham, MA, USA) and by gel electrophoresis, respectively. DNA of sample was stored at -20°C for further study. Haematoxylin Eosin (H&E) staining were prepared according to the method of [12, 13]. Samples were fixed in 10% buffered formalin, and embedded in paraffin. Three to five micrometer thick sections were stained with hematoxylin (Sigma H 3136) for 10 min and with eosin (Sigma E 4382) for 1 min to establish the diagnosis areas. The Research Ethics Committee of General Hospital of PLA approved the collection of tissue samples for research. Written informed consent was obtained from individuals in order to collect the tissues during surgery.

### *DNA methylation methods*

Bisulfite conversion of 500 ng genomic DNA was performed using the EZ DNA methylation kit (Zymo Research). DNA methylation level was assessed according to the manufacturer's instructions using Infinium-HumanMethylation450 Beadchips (Illumina Inc.). The technical schemes, the accuracy, and the high reproducibility of this array have been described in previous papers [14]. Quantitative measurements of DNA methylation were determined for 485,577 CpG dinucleotides, which covered 99% of the RefSeq genes and were distributed across the whole gene regions, including promoter, gene body, and 30-untranslated regions (UTRs). They also covered 96% of CGIs from the UCSC database with additional coverage in CGI shores (0-2 kb from CGI) and CGI shelves (2-4 kb from CGI). Detailed information on the contents of the array is available in the Infinium HumanMethylation450 User Guide and HumanMethylation450 manifest (www.illumina.com) and in recent papers [15]. DNA methylation data were analyzed with the methylation analysis module within the BeadStudio software (Illumina Inc.). DNA methylation status of the CpG sites was calculated as the ratio of the signal from a methylated probe relative to the sum of both methylated and unmethylated probes. This value, known as b, ranges from 0 (completely unmethylated) to 1 (fully methylated). For intra-chip normalization of probe intensities, colored balance and background corrections in every set of ten samples from the same

**Table 1.** Haematoxylin and eosin (H&E) staining analysis of pancreatic cancer tissue and tissue adjacent to cancer of six patients

| Patients | Age | Haematoxylin Eosin (H&E) staining analysis | |
|---|---|---|---|
| | | B (Pericarcinous tissue) | C (Pancreatic cancer tissue) |
| Patient 1 (F) | 74 | B1 (&) | C1 (&) |
| Patient 2 (M) | 36 | B2 (*) | C2 (&) |
| Patient 3 (M) | 66 | B3 (&) | C3 (&) |
| Patient 4 (F) | 60 | B4 (*) | C4 (&) |
| Patient 5 (F) | 46 | B5 (&) | C5 (&) |
| Patient 6 (F) | 71 | B6 (&) | C6 (&) |

Note: F means female, M means male. &represents qualified sample, *represents unqualified samples.

chip were performed using internal control probes. X chromosome CpG sites in the CGIs in the AR gene in this array as well as the internal control probes were checked to validate the DNA methylation measurements.

*Statistical methods*

Surrogate variable analysis was used to identify CpG loci showing significant differences in DNA methylation between pancreatic cancer tissues and pericarcinous tissues [16]. This analysis is useful in clinical studies, where a large number of clinical variables, including known and unknown factors, have a complicated joint impact on microarray data, as applied in previous studies [17]. A false discovery rate (FDR) correction was applied at the 0.05 level for multiple testing. A paired t-test was used to assess the significance of DNA methylation differences between pancreatic cancer tissue and tissue adjacent to cancer. $p$ values <0.01 were considered significant differential methylation of DNA methylation differences between two groups.

*Bioinformatics*

The biological processes, molecular functions, and cellular components of the identified differentially methylated genes were examined by using the KOBAS software to perform gene ontology (GO) annotation (http://kobas.cbi.pku.edu.cn/home.do) [18]. GO enrichment analysis of differentially methylated genes was implemented by GORILLA (http://cbl-gorilla.cs.technion.ac.il/) [19], in which gene length bias was corrected. GO terms with corrected $p$-value less than $10^{-4}$ were considered significantly enriched by differential methylated genes. KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-through put experimental technologies (http://www.genome.jp/kegg/). We used KOBAS software to test the statistical enrichment of differential methylated genes in KEGG pathways.

## Results

*Histopathologic features*

For each patient, one pancreatic cancer tissue and one pericarcinous tissue were collected respectively. Twelve tissues, including six pancreatic cancer tissues and six pericarcinous tissues, were gathered from six patients according our rigorous set of criteria as described in method section (**Table 1**). Then all of the tissues were implemented haematoxylin and eosin (H&E) staining to test the accurate rate of sampling, the results indicated that pericarcinous tissues of patients 2 and patients 4 cannot reach the standers that we set were excluded from this study. **Figure 1** showed two examples of pancreatic cancer tissues and two pericarcinous tissues after Haematoxylin Eosin (H&E) staining using optical microscope with different magnification. The islands of tumor cells were surrounded by an abundant mucinous stroma with separation of the tumor cells from the surrounding stroma (clefting). Prominent increased vascularity was also noted, as well as solar elastosis and variable amounts of predominantly mononuclear inflammatory cells with scattered neutrophils.

*Diagnostic differences in DNA methylation between pancreatic cancer tissues with pericarcinous tissues*

DNA methylation levels were compared between six pancreatic cancer tissues with four pericarcinous tissues control subjects using Infinium HumanMethylation450 Bead Chips. Of 485,577 CpG sites, significant diagnostic differences in DNA methylation were observed at 24,417 CpG sites representing 9,589 genes at

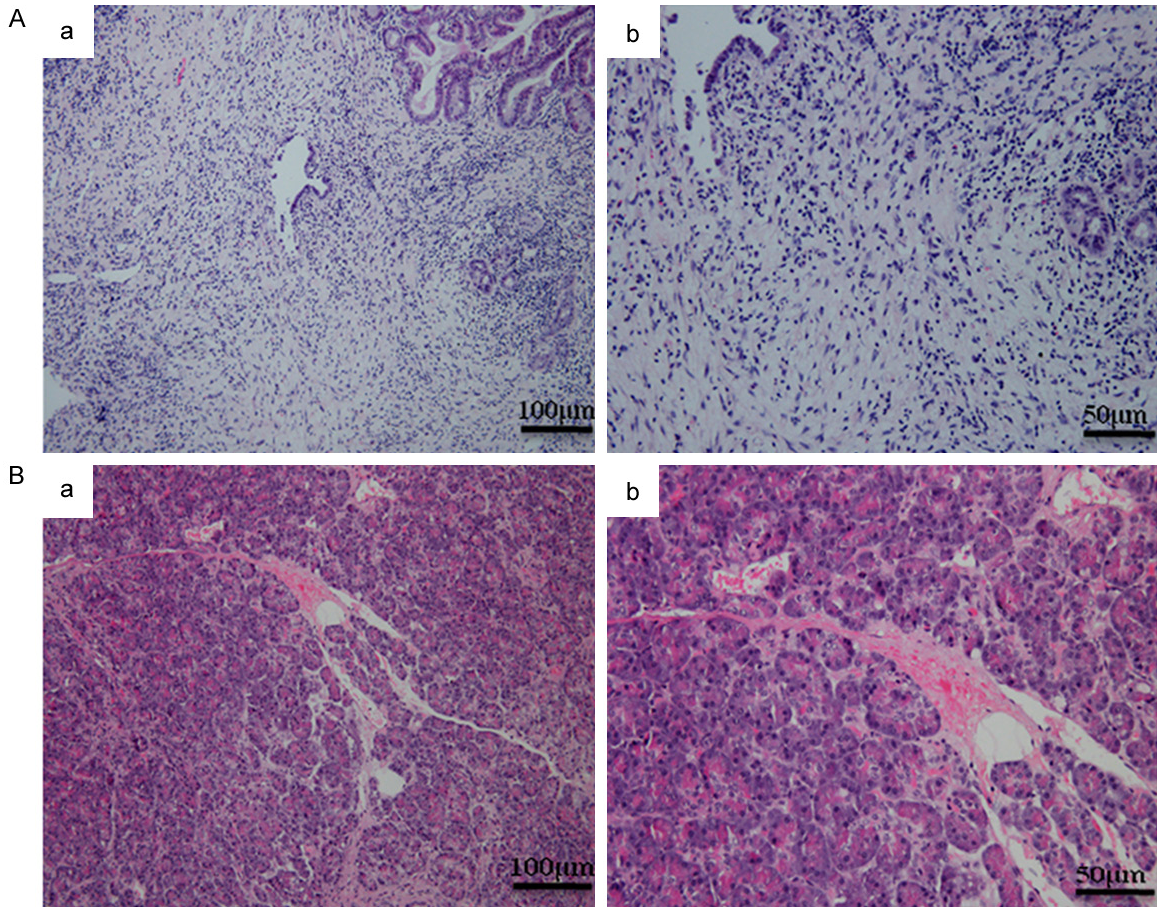


**Figure 1.** Aa. Pericarcinous tissue with 100 times magnification. Ab. Pericarcinous tissue with 200 times magnification. Ba. Pancreatic cancer tissue with 100 times magnification. Bb. Pancreatic cancer tissue with 200 times magnification.

FDR 5% correction (**Figure 2**). Of 24,417 CpG sites with significant diagnostic differences in DNA methylation, 14,721 (60%, 14,721 of 24,417) CpG sites were hypomethylated and 9,705 (40%, 9,705 of 24,417) CpG sites were hypermethylated. Functional distribution of 9,705 hypermethylated CpG sites suggested that 42% of these sites were located in promoter regions, 37% of these sites were located in gene bodies, 16% of these sites were located in intergenic regions and 5% of these sites were located in the 3'-untranslated regions (UTRs). Furthermore, sublocation analysis of 4032 CpG sites in promoter region with hypermethylated indicated that 35% of these sites were located in regions from -200 to -1,500 nt upstream of the transcription start site (TSS1500), 26% of these sites were located in regions from -200 nt upstream to the TSS itself (TSS200), 24% of these sites were located in 1st Exon regions and 15% of these sites were located in the 5'-untranslated regions (UTRs). These hypermethylated CpG sites were mostly located in gene bodies and promoter regions. Meanwhile, Functional distribution of 14,721 hypomethylated CpG sites suggested that 22% of these sites were located in promoter regions, 39% of these sites were located in gene bodies, 35% of these sites were located in intergenic regions and 4% of these sites were located in 3'UTR regions. Furthermore, sublocation analysis of 14,721 hypomethylated CpG sites in promoter regions indicated that 50% of these sites were located in TSS1500 regions, 16% of these sites were located in TSS200 regions, 13% of these sites were located in 1st Exon regions and 21% of these sites were located in 5'UTR regions. These hypomethylated CpG sites were mostly located in gene bodies, promoter regions and intergenic regions. The results above seem to be conflictive with the previous studies, which showed that gene promoter areas of actively

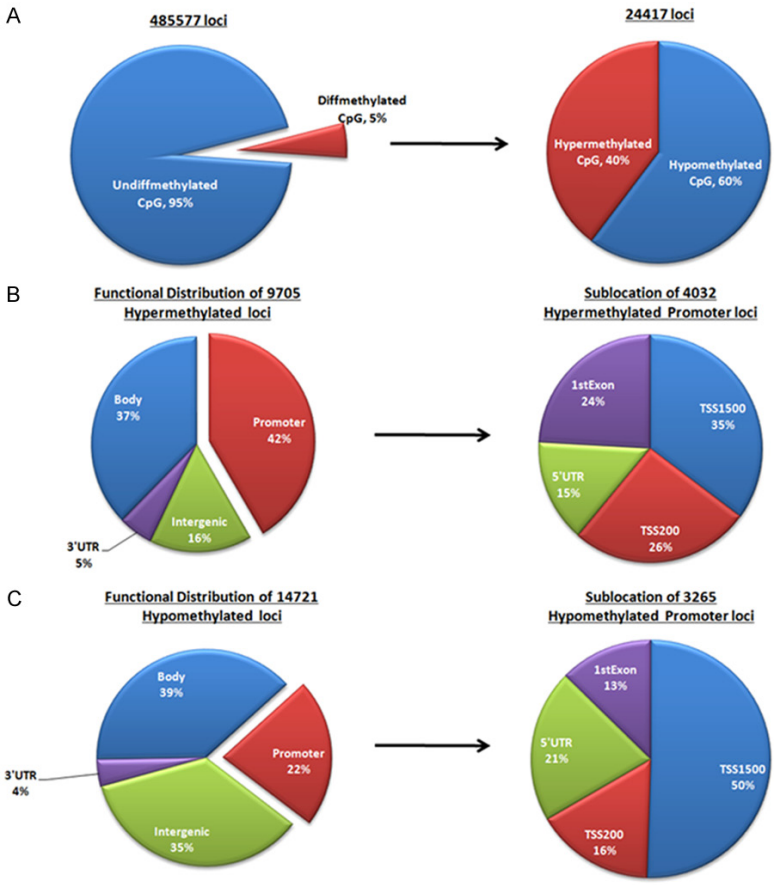


**Figure 2.** Graphic illustration of functional distribution and differential methylated CpG sites identified in this study.

transcribed genes are largely unmethylated so as to be accessible to transcription factors. Based on the research object, we speculated that lots of transcription factor must be modified and detained from transcription so that they can adapt to the cellular abnormalities as cancer.

Because the 24,417 methylated CpG sites corresponded to 9,589 genes, some of the methylated genes must contain more than one methylated site. Further analysis showed that among the 9,589 methylated genes, 5685 (59%) possessed only one methylated site, 2133 (22%) contained two methylated sites, and 1771 (18%) contained three or more methylated sites (**Figure 3**). In particular, one methylated gene (PTPRN2, ENSG00000155093) with protein tyrosine phosphatase possessed 73 methylated sites. The MAD1 mitotic arrest deficient-like 1 (yeast) (MAD1L1, ENSG00000002822), adenosine deaminase (ADARB2, ENSG0000018-5736) and tenascin XB (TNXB, ENSG00000231608) all possessed over 30 methylated sites.

*Gene ontology (GO) and KEGG pathway analysis of differential methylated genes*

In order to improve the credibility of this research, the genes with counts of methylation sites were equal or greater than three were selected to perform intensive study. After such screening, 1771 genes with more than three counts of differential methylated CpG sites were harvested. GO terms were further assigned to Homo sapiens differential methylated genes based on their sequence similarities to known proteins in the UniProt database annotated with GO terms as well as InterPro and Pfam domains they contain. GO annotation of all 1771 differentially methylated genes was implemented by KOBAS software (http://kobas.cbi.pku.edu.cn/home.do) and GORILLA (http://cbl-gorilla.cs.technion.ac.il/) was performed to GO enrichment, in which gene length bias was corrected. GO terms with corrected *p*-value less than $10^{-4}$ were considered significantly enriched (**Figure 4**). Biological processes, cellular components, and molecular functions are shown in **Figure 4**. From the perspective of biological processes, there are 430 GO terms were assigned under this catalogues. Among these terms, homophilic cell adhesion via plasma membrane adhesion molecules (GO: 0007156, *p* value: 3.65E-13), cell morphogenesis involved in differentiation (GO: 0000904, *p* value: 1.29E-06) and cell adhesion (GO: 0007155, *p* value: 1.07E-06) were the top three significantly enriched terms. From the cellular component perspective, there are 103 GO terms were assigned under this catalogues. Among these terms, cell projection (GO: 0042995, *p* value: 5.75E-07) was the top significantly enriched terms. From the molecu-

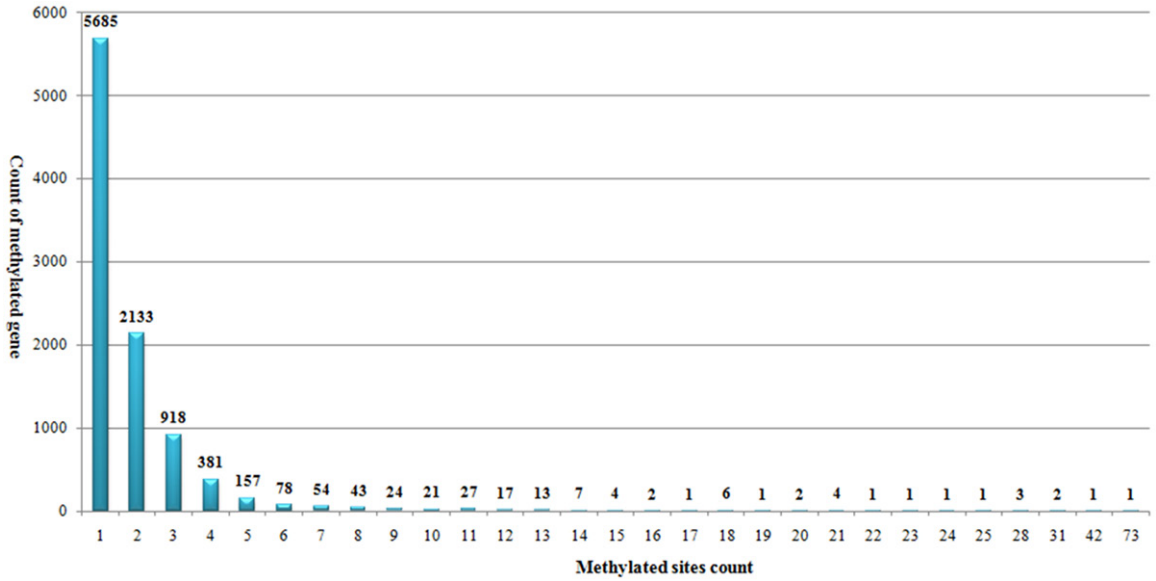


**Figure 3.** Analysis of the identified methylated CpG sites. Distribution of the methylated CpG sites in the methylated genes.

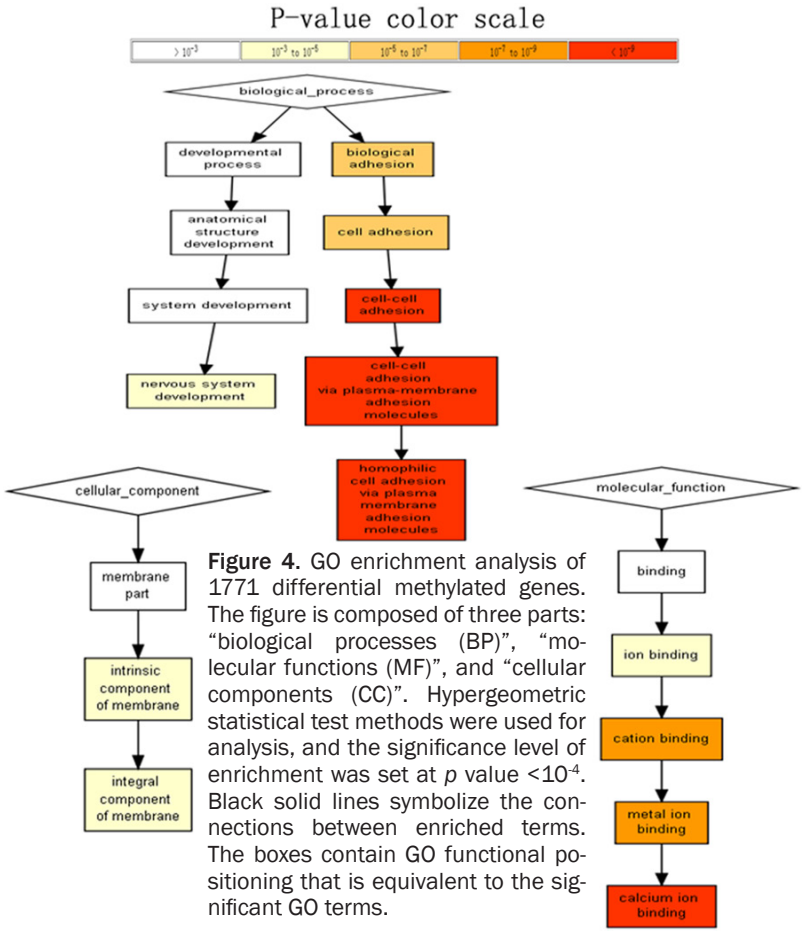


**Figure 4.** GO enrichment analysis of 1771 differential methylated genes. The figure is composed of three parts: "biological processes (BP)", "molecular functions (MF)", and "cellular components (CC)". Hypergeometric statistical test methods were used for analysis, and the significance level of enrichment was set at $p$ value $<10^{-4}$. Black solid lines symbolize the connections between enriched terms. The boxes contain GO functional positioning that is equivalent to the significant GO terms.

lar function perspective, there are 135 GO terms were assigned under this catalogues. Among these terms, calcium ion binding (GO: 0005509, $p$ value: 4.63E-06) and were the top significantly over-represented terms.

In vivo, various biological functions were implemented by cooperation of different genes. Pathways enrichment analysis can give some clues to the biochemical and signal transduction pathways that differential expression genes may participate in. KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular level information, especially large-scale molecular datasets generated by genome sequencing and other high throughput experimental technologies (http://www.genome.jp/kegg/). We used KOBAS software to test the statistical enrichment of differential methylated genes in KEGG path-

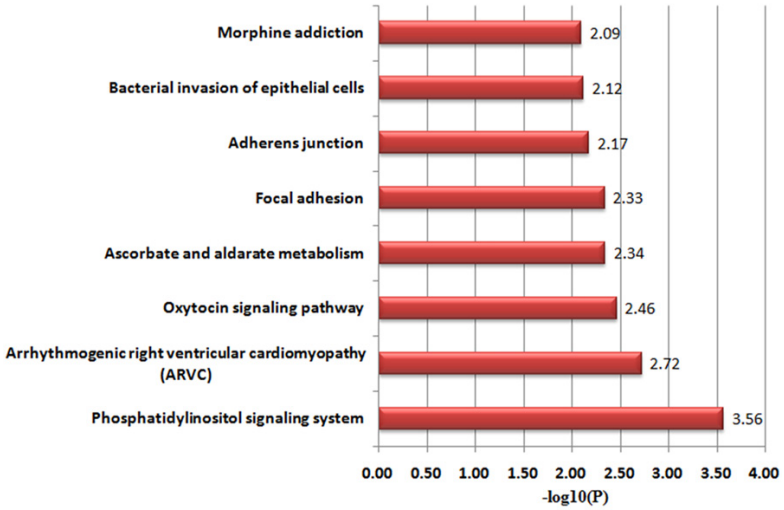


**Figure 5.** KEGG enrichment analysis of all the 1771 differential methylated genes identified in this study. The significance level of enrichment was set at *p* value <0.01.

ways [20]. In this study, 1771 differential methylated genes involve 261 pathways. **Figure 5** shows the results of pathways enrichment, it clearly display that phosphatidylinositol signaling system were the top enriched term. 22 differential methylated genes that identified in our study participate in this pathway. Moreover, it is worth noting that arrhythmogenic right ventricular cardiomyopathy (ARVC), oxytocin signaling pathway, ascorbate and aldarate metabolism and Focal adhesion were also significant enriched in this study. The pathways mentioned above were adopted with the function that cancer played.

*Candidate genes regulated by aberrant DNA methylation*

To pinpoint the candidate genes regulated by aberrant DNA methylation in pancreatic cancer, we focused on the genes of discordant categories in **Figure 6**. Genes with count of methylation was equal or greater than 13 were selected to perform further study (*p* value <0.05). After such screening, 51 genes with equal or more than 13 counts of methylated sites were harvested. Functional descriptions of these genes suggested they owned a wide range. For example, PTPRN2 gene as the most significant methylated item possessed 73 methylated CpG sites, which was identified as protein tyrosine phosphatase, MAD1L1 gene as the second significant methylated item possessed 42 methylated CpG sites, which was identified as MAD1 mitotic arrest deficient-like 1. GO analysis of these 51 genes revealed function positioning from three different viewpoints (**Figure 7**). From the perspective of biological processes, homophilic cell adhesion via plasma membrane adhesion molecules (GO: 0007156, *p* value: 1.27E-08), cell-cell adhesion via plasma-membrane adhesion molecules (GO: 0098742, *p* value: 1.34E-07) and nervous system development (GO: 0007399, *p* value: 0.026) were the top three significantly enriched terms. From the cellular component perspective, intrinsic component of membrane (GO: 0031224, *p* value: 0.063) was the top significantly enriched terms. From the molecular function perspective, calcium ion binding (GO: 0005509, *p* value: 0.00011) and protein kinase A catalytic subunit binding (GO: 0034236, *p* value: 0.038) was the top two significantly enriched terms. Meanwhile, KEGG analysis indicated that Insulin signaling pathway (hsa04910, *p* value: 0.01), Circadian entrainment (hsa04713, *p* value: 0.036) and Platelet activation (hsa04611, *p* value: 0.06) were the most enriched in this study.

## Discussion

Feinberg and Vogelstein were the first to associate differences in DNA methylation status to cancer in the early 80' of the previous century [21]. Since then, it has been an explosion of research regarding aberrant DNA methylation of various diseases. Recently developed genome scale methods for mapping DNA methylation across whole genomes have resulted in translation of basic discoveries. Therefore, aberrant methylation of genes which were associated with diseases, including types of cancer, Prader-Willi, Angelman and so on, were introduced into routine clinical application.

In this study, genome-wide DNA methylation profiling was conducted in pancreatic cancer tissues and pericarcinous tissues (six pancreatic cancer tissues and four pericarcinous tissues) using Infinium HumanMethylation450

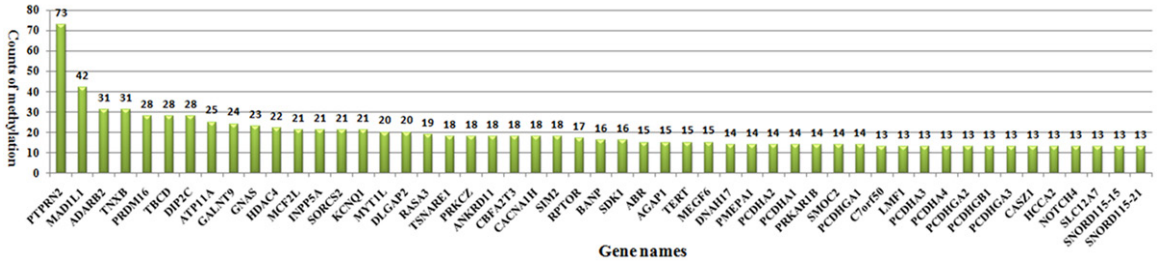


**Figure 6.** 51 candidate genes with aberrant DNA methylation.

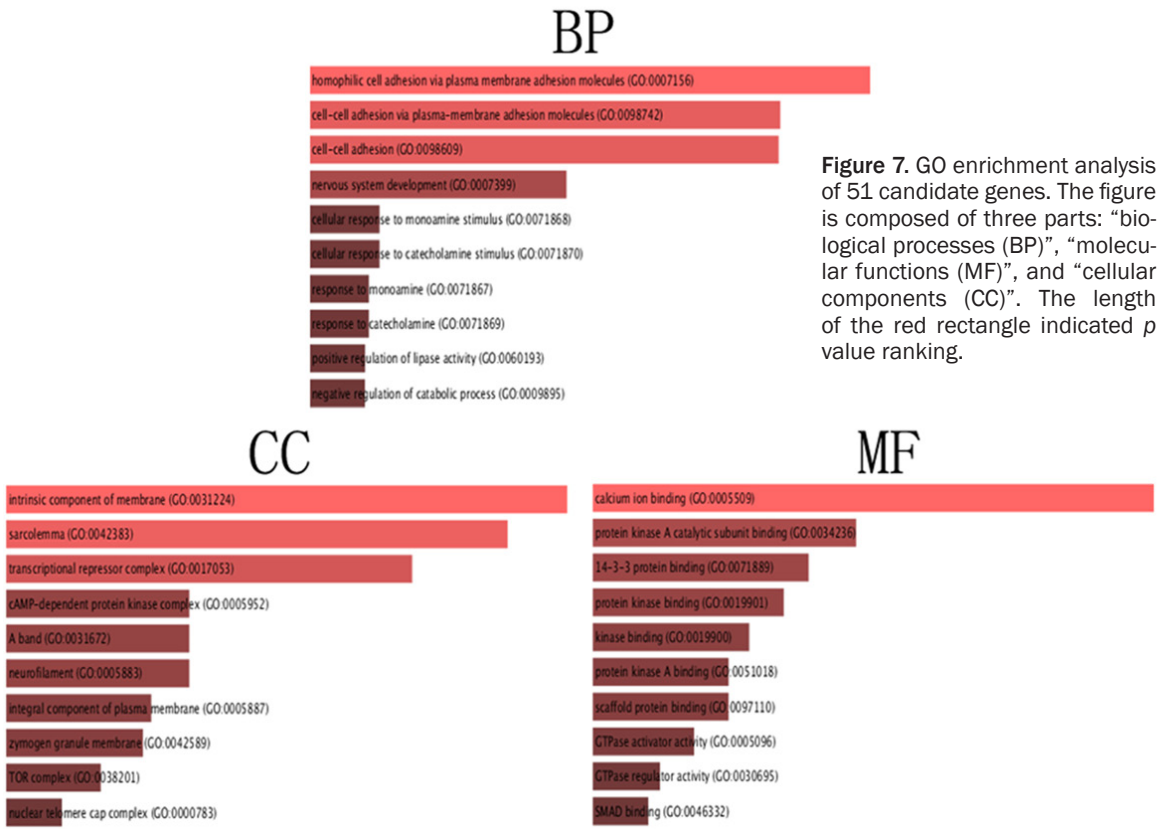


**Figure 7.** GO enrichment analysis of 51 candidate genes. The figure is composed of three parts: "biological processes (BP)", "molecular functions (MF)", and "cellular components (CC)". The length of the red rectangle indicated *p* value ranking.

Beadchips. DNA methylation is associated with genotypic variants [22]; sampling from pancreatic cancer tissues and pericarcinous tissues of one patient is a useful method for investigating DNA methylation differences between disease phenotypes without the influence of genetic discordance. In fact, this approach has been applied successfully to identify epigenetic differences in non-small cell lung cancer [23], colorectal carcinoma [24] and hepatocellular carcinoma [25]. In this study, a total of 24,417 differentially methylated CpG sites (5%, of 485,577 CpG sites,) corresponding 9,589 genes that were identified between the two cohorts. Of 24,417 CpG sites with significant diagnostic differences in DNA methylation, 14,721 (60%, 14,721 of 24,417) CpG sites were hypomethylated and 9,705 (40%, 9,705 of 24,417) CpG sites were hypermethylated. Functional distribution of 9,705 hypermethylated CpG sites suggested that 42% of these sites were located in promoter regions, 37% of these sites were located in gene bodies. Meanwhile, Functional distribution of 14,721 hypomethylated CpG sites suggested that 22% of these sites were located in promoter regions, 39% of these sites were located in gene bodies. The results mentioned above indicated

that hypomethylated CpG sites seem to be more prevalent in the whole genome. On the contrary, hypermethylated CpG sites can be more easily indentified in promoter regions. As shown in **Figure 2**, aberrant DNA methylation in pancreatic cancer was mostly observed at CpG sites in the gene bodies and promoter regions. The present study demonstrated that altered DNA methylation in pancreatic cancer tissues occurred at CpG sites not only in the CGIs but also in CGI shores and CGI shelves. This is consistent with previous research that is methylation of these regions can represses transcription. For example, Irizarry et al. demonstrated that altered DNA methylation in cancer occurred in CGI shores rather than in the CGIs, and DNA methylation changes in CGI shores were strongly related to gene expression [26]. In addition, we had noticed that numerous differential CpG sites were located in gene bodies. It is still unknown about the mechanism that how these differential CpG sites can have impact on gene functions. Shann et al. demonstrated the correlation between intragenic hypomethylation and gene silencing in cancer cell lines [27], and Ball et al. demonstrated that gene body DNA methylation in highly expressed genes is a consistent phenomenon in human cells [28]. Recently, it became apparent that CGIs in gene bodies act as alternative promoters [29, 30] and that tissue-specific or cell type-specific CGI methylation is prevalent in gene bodies [30, 31]. GO enrichment analysis suggested that significant function differences were identified between pancreatic cancer tissues with pericarcinous tissues. Moreover, KEGG analysis indicated several important pathways can be retrieved in this study. For example, previous studies suggested that phosphatidylinositol 3-kinases in the most enriched item of phosphatidylinositol signaling system (hsa04070) play important role PI3K/Akt/mTOR signaling pathway, which has impact on proliferation and activation of tumor cell [32, 33], and this pathway is becoming to research highlights of targeted drug. Moreover, we have noticed that focal adhesion kinase (FAK) which belongs to Focal adhesion (hsa04510) pathway. FAK is a focal adhesion-associated protein kinase involved in cellular adhesion (how cells stick to each other and their surroundings) and spreading processes (how cells move around) [34]. It has been shown that when FAK was blocked, breast cancer cells became less metastatic due to decreased mobility [35]. FAK is phosphorylated in response to integrin engagement, growth factor stimulation, and the action of mitogenic neuropeptides [36, 37].

There are several limitations to the present study. First, the sample size was not large. Replication studies will be needed in larger samples. Second, the analyzed CpG sites were limited in number, although the 450 K microarray is one of the most powerful and cost-effective tools currently available for assessing methylation changes. Third, it is not possible to differentiate methylation from 5-hydroxymethylation of cytosine, which also plays a critical role in gene regulation [38]. In summary, aberrant DNA methylation in pancreatic cancer tissues was identified at numerous CpG sites across the whole genome in using two independent sets of samples. Of the differently methylated CpG sites in the CGIs, most of them were located in the promoter regions. These findings support the hypothesis that altered DNA methylation could be involved in the pathophysiology of pancreatic cancer. In this study, we analyzed the genome-wide DNA methylation profiles of human somatic tissues. Although the number of analyzed individuals was limited, the analysis was sufficient to provide DNA methylation distribution patterns across different genomic regions that were largely in agreement with patterns previously observed by previous studies. The methylome data alone was sufficient for correctly distinguishing between all the ten tissues studied, collectively demonstrating that tissues are characterized by distinctive methylation patterns that reflect their tissue-specific functions. Our study provoked the question, of how differential methylated CpG sites mechanistically contribute to the gene functions, especially for the numerous methylation regions that were found in gene body areas. In addition, it remains unclear, however, how the gene body differential methylated CpG sites may function as regulators of gene expression, and this question should be addressed in the future epigenetic studies.

## Conclusion

Previous studies have demonstrated that DNA methylation play important roles in the regulation of developmental processes of several types cancers. In this study, a total of 24,417

CpG sites representing 9,589 genes that were common between these two cohorts were identified. Of the 24,417 CpG sites, 14,721 (60%, 14,721 of 24,417) CpG sites were hypomethylated and 9,705 (40%, 9,705 of 24,417) CpG sites were hypermethylated. Then GO (Gene Ontology) and KEGG analysis were implemented to systematically characterize the significant differential methylated genes between pancreatic cancer tissues and pericarcinous tissues. In addition, we further screened 51 genes with aberrant methylation, which were the most likely candidate methylation markers within the scale of global differential methylation profiling. GO and KEGG analysis indicated these genes owning a wide range of functions. The identification of differential methylated genes in this study provides information valuable to the in-depth study of pancreatic cancer. Moreover, the results of this study will not only further our understanding of the differential methylated genes, but will also help to enhance methylome studies of pancreatic cancer.

## Acknowledgements

This study was supported by the National Natural Science Foundation of China (No: 811-70494), Natural Science Foundation of Beijing (No: 7162176) and Beijing Nova program (Z121107002512122).

## Disclosure of conflict of interest

None.

**Address correspondence to:** Dr. Qingfang Li, Department of Cancer Center, General Hospital of PLA, 28 Fuxing Road, Beijing 100853, China. Tel: +86-10-66937592; Fax: +86-10-66937592; E-mail: liqfluck@126.com

## References

[1] Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ. Cancer statistics, 2008. CA Cancer J Clin 2008; 58: 71-96.

[2] Li D, Xie K, Wolff R, Abbruzzese JL. Pancreatic cancer. Lancet 2004; 363: 1049-57.

[3] Hassan MM, Bondy ML, Wolff RA, Abbruzzese JL, Vauthey JN, Pisters PW, Evans DB, Khan R, Chou TH, Lenzi R, Jiao L, Li D. Risk factors for pancreatic cancer: case-control study. Am J Gastroenterol 2007; 102: 2696-707.

[4] Batty GD, Kivimaki M, Morrison D, Huxley R, Smith GD, Clarke R, Marmot MG, Shipley MJ. Risk factors for pancreatic cancer mortality: extended follow-up of the original Whitehall Study. Cancer Epidemiol Biomarkers Prev 2009; 18: 673-5.

[5] Landi S. Genetic predisposition and environmental risk factors to pancreatic cancer: A review of the literature. Mutat Res 2009; 681: 299-307.

[6] Wolpin BM, Chan AT, Hartge P, Chanock SJ, Kraft P, Hunter DJ, Giovannucci EL, Fuchs CS. ABO blood group and the risk of pancreatic cancer. J Natl Cancer Inst 2009; 101: 424-31.

[7] Tersmette AC, Petersen GM, Offerhaus GJ, Falatko FC, Brune KA, Goggins M, Rozenblum E, Wilentz RE, Yeo CJ, Cameron JL, Kern SE, Hruban RH. Increased risk of incident pancreatic cancer among first-degree relatives of patients with familial pancreatic cancer. Clin Cancer Res 2001; 7: 738-44.

[8] Petronis A, Gottesman II, Crow TJ, DeLisi LE, Klar AJ, Macciardi F, McInnis MG, McMahon FJ, Paterson AD, Skuse D, Sutherland GR. Psychiatric epigenetics: a new focus for the new century. Mol Psychiatry 2000; 5: 342-6.

[9] Alvarez H, Opalinska J, Zhou L, Sohal D, Fazzari MJ, Yu Y, Montagna C, Montgomery EA, Canto M, Dunbar KB, Wang J, Roa JC, Mo Y, Bhagat T, Ramesh KH, Cannizzaro L, Mollenhauer J, Thompson RF, Suzuki M, Meltzer SJ, Melnick A, Greally JM, Maitra A, Verma A. Widespread hypomethylation occurs early and synergizes with gene amplification during esophageal carcinogenesis. PLoS Genet 2011; 7: e1001356.

[10] Vincent A, Omura N, Hong SM, Jaffe A, Eshleman J, Goggins M. Genome-wide analysis of promoter methylation associated with gene expression profile in pancreatic adenocarcinoma. Clin Cancer Res 2011; 17: 4341-54.

[11] Nones K, Waddell N, Song S, Patch AM, Miller D, Johns A, Wu J, Kassahn KS, Wood D, Bailey P, Fink L, Manning S, Christ AN, Nourse C, Kazakoff S, Taylor D, Leonard C, Chang DK, Jones MD, Thomas M, Watson C, Pinese M, Cowley M, Rooman I, Pajic M; APGI, Butturini G, Malpaga A, Corbo V, Crippa S, Falconi M, Zamboni G, Castelli P, Lawlor RT, Gill AJ, Scarpa A, Pearson JV, Biankin AV, Grimmond SM. Genome-wide DNA methylation patterns in pancreatic ductal adenocarcinoma reveal epigenetic deregulation of SLIT-ROBO, ITGA2 and MET signaling. Int J Cancer 2014; 135: 1110-8.

[12] Chan JK. The wonderful colors of the hematoxylin-eosin stain in diagnostic surgical pathology. Int J Surg Pathol 2014; 22: 12-32.

[13] Fischer AH, Jacobson KA, Rose J, Zeller R. Hematoxylin and eosin staining of tissue and cell sections. CSH Protoc 2008; 2008: pdb.prot 4986.

[14] Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Epigenomics 2011; 3: 771-84.
[15] Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, Esteller M. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics 2011; 6: 692-702.
[16] Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet 2007; 3: 1724-35.
[17] Numata S, Ye T, Hyde TM, Guitart-Navarro X, Tao R, Wininger M, Colantuoni C, Weinberger DR, Kleinman JE, Lipska BK. DNA methylation signatures in development and aging of the human prefrontal cortex. Am J Hum Genet 2012; 90: 260-72.
[18] Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY, Wei L. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res 2011; 39: W316-22.
[19] Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 2009; 10: 48.
[20] Mao X, Cai T, Olyarchuk JG, Wei L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics 2005; 21: 3787-3793.
[21] Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 1983; 301: 89-92.
[22] Numata S, Ye T, Hyde TM, Guitart-Navarro X, Tao R, Wininger M, Colantuoni C, Weinberger DR, Kleinman JE, Lipska BK. DNA methylation signatures in development and aging of the human prefrontal cortex. Am J Hum Genet 2012; 90: 260-72.
[23] Wang LG, Ni Y, Su BH, Mu XR, Shen HC, Du JJ. MicroRNA-34b functions as a tumor suppressor and acts as a nodal point in the feedback loop with Met. Int J Oncol 2013; 42: 957-62.
[24] Jin S, Mu Y, Wang X, Liu Z, Wan L, Xiong Y, Zhang Y, Zhou L, Li L. Overexpressed RACK1 is positively correlated with malignant degree of human colorectal carcinoma. Mol Biol Rep 2014; 41: 3393-9.
[25] Wang Z, Zhang L, Zhang D, Sun R, Wang Q, Liu X. Glycolysis inhibitor 2-deoxy-D-glucose suppresses carcinogen-induced rat hepatocarcinogenesis by restricting cancer cell metabolism. Mol Med Rep 2015; 11: 1917-24.
[26] Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 2009; 41: 178-86.
[27] Shann YJ, Cheng C, Chiao CH, Chen DT, Li PH, Hsu MT. Genome-wide mapping and characterization of hypomethylated sites in human tissues and breast cancer cell lines. Genome Res 2008; 18: 791-801.
[28] Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 2009; 27: 361-8.
[29] Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr AR, James KD, Turner DJ, Smith C, Harrison DJ, Andrews R, Bird AP. Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet 2010; 6: e1001134.
[30] Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, Turecki G, Delaney A, Varhol R, Thiessen N, Shchors K, Heine VM, Rowitch DH, Xing X, Fiore C, Schillebeeckx M, Jones SJ, Haussler D, Marra MA, Hirst M, Wang T, Costello JF. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 2010; 466: 253-7.
[31] Deaton AM, Webb S, Kerr AR, Illingworth RS, Guy J, Andrews R, Bird A. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. Genome Res 2011; 21: 1074-86.
[32] Yu G, Huang B, Chen G, Mi Y. Phosphatidylethanolamine-binding protein 4 promotes lung cancer cells proliferation and invasion via PI3K/Akt/mTOR axis. J Thorac Dis 2015; 7: 1806-1816.
[33] Tang H, Li RP, Liang P, Zhou YL, Wang GW. miR-125a inhibits the migration and invasion of liver cancer cells via suppression of the PI3K/AKT/mTOR signaling pathway. Oncol Lett 2015; 10: 681-686.
[34] Blackshaw SE, Kamal DJ, Lackie JM. The dictionary of cell and molecular biology. 3rd edition. San Diego: Academic Press; 1999. ISBN 0-12-432565-3.
[35] Leslie M. Physorg: When cancer cells can't let go. J Cell Biol 2009; 185: 178.
[36] Chen R, Kim O, Li M, Xiong X, Guan JL, Kung HJ, Chen H, Shimizu Y, Qiu Y. Regulation of the PH-domain-containing tyrosine kinase Etk by focal adhesion kinase through the FERM domain. Nat Cell Biol 2001; 3: 439-44.
[37] Eliceiri BP, Puente XS, Hood JD, Stupack DG, Schlaepfer DD, Huang XZ, Sheppard D, Cheresh DA. Src-mediated coupling of focal adhesion kinase to integrin alpha (v) beta5 in vascular endothelial growth factor signaling. J Cell Biol 2002; 157: 149-60.
[38] Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. Cell 2011; 146: 866-72.