Journal of Bone Oncology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Bone Oncology

journal homepage: www.elsevier.com/locate/jbo

Research Paper

An integrative analysis of DNA methylation in osteosarcoma

Jie Xu, Deng Li, Zhiqing Cai, Yingbin Zhang, Yulin Huang, Baohua Su, Ruofan Ma*

Department of Orthopedics, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou 510120, China

ARTICLE INFO

Keywords:
Differentially methylated gene
Osteosarcoma
Protein-protein interaction
Transcription factor

ABSTRACT

Background: The study aimed to analyze aberrantly methylated genes, relevant pathways and transcription factors (TFs) in osteosarcoma (OS) development.

Methods: Based on the DNA methylation microarray data GSE36002 that were downloaded from GEO database, the differentially methylated genes in promoter regions were identified between OS and normal samples. Pathway and function enrichment analyses of differentially methylated genes was performed. Subsequently, protein-protein interaction (PPI) network was constructed, followed by identification of cancer-associated differentially methylated genes and significant differentially methylated TFs.

Results: A total of 1379 hyper-methylation regions and 169 hypo-methylation regions in promoter regions were identified in OS samples compared to normal samples. The differentially hyper-methylated genes were significantly enriched in Neuroactive ligand-receptor interaction pathway, and Peroxisome proliferator activated receptor (PPAR) signaling pathway. The differentially hypo-methylated genes were significantly enriched in Toll-like receptor signaling pathway. In PPI network, signal transducers and activators of transcription (STAT3) had high degree (degree = 21). MAX interactor 1, dimerization protein (MXI1), STAT3 and T-cell acute lymphocytic leukemia 1 (TAL1) were significant TFs enriched with target genes in OS samples. They were found to be cancer-associated and hyper-methylated in OS samples.

Conclusion: Neuroactive ligand-receptor interaction, PPAR signaling, Toll-like receptor signaling pathways are implicated in OS. MXI1, STAT3, and TAL1 may be important TFs involved in OS development.

1. Introduction

Osteosarcoma (OS) is the most common primary malignant bone tumor in children and adolescents. More than 15% of patients with OS develop metastases, frequently to lung [1]. For patients with metastasis or recurrence, the long-term survival rate is less than 20% [2,3]. Understanding of the molecular mechanisms of OS would provide a basis for developing new therapeutic strategies.

It has been demonstrated that genetic alternations in the status of DNA methylation count among the most common molecular alterations in human neoplasia [4]. DNA methylation usually results in obstruction of the promoter region, hampering gene transcription and causing gene silencing [5]. Quite a few studies have reported findings related to DNA methylation in OS. It has been reported that methylation of frizzled-related proteins (SFRPs) may promote Wnt signaling pathway, thereby enhancing OS cell invasion [6]. Hyper-methylation of p14ADP-ribosylation factor (ARF) and estrogen receptor 1 (ESR1) have been found in OS as well, and may have implications in the prognosis of OS patients [7]. Moreover, Lu et al. [8] have reported that iroquois homeobox 1 (IRX1) enhances OS metastasis and may be a potential molecular marker. Recent study [9] suggests that promoter hyper-methylation of reversion-inducing cysteine-rich protein with Kazal motifs (RECK) is a causative factor in metastasis of OS. Despite some advances have been achieved in this field, certain mechanisms underlying OS remain largely unknown. Microarray analysis has been applied to identify the gene alterations and screen potential targets in human OS cell lines [10]. Kresse et al. [11] integrated genomewide genetic and epigenetic profiles from the EuroBoNeT panel, a European Network of Excellence on bone tumors (http://www.eurobonet.eu), of 19 human osteosarcoma cell lines based on microarray technologies, and deposited the DNA methylation dataset in the Gene Expression Omnibus (GEO) data repository (accession number GSE36002). In their study, they have comprehensively analyzed the relationships of DNA copy number, DNA methylation and mRNA expression in OS. Additionally, they screened out the differentially methylated genes and performed functional enrichment analysis. However, the interaction between differentially methylated genes, and the classification of these genes have not been analyzed.

Since methylation of CpG islands in promoter regions is a mechanism for inactivating genes, we estimated the differentially

* Corresponding author.
E-mail address: maruofan_syx@163.com (R. Ma).

http://dx.doi.org/10.1016/j.jbo.2017.05.001
Received 19 January 2017; Received in revised form 9 May 2017; Accepted 12 May 2017

methylated genes in promoter regions between OS and normal samples from GSE36002 in this study. Pathway and functional enrichment analyses of differentially methylated genes was then performed. Subsequently, protein-protein interaction (PPI) network was constructed to analyze the interactions between differentially methylated genes. Furthermore, the classifications of differentially methylated genes, including tumor suppressor (TS) genes, oncogenes and TFs were investigated.

2. Materials and methods

2.1. DNA methylation microarray data

The GSE36002 DNA methylation microarray data [11] were downloaded from GEO database in National Center of Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/gds/), based on the platform of Illumina Human Methylation 27 BeadChip (Illumina Inc., California, USA). GSE36002 dataset is comprised of 25 samples: 19 osteosarcoma cell lines samples and 6 normal control samples (four normal bone samples and two osteoblast cultures).

2.2. Identification of differentially methylated genes

DNA methylation microarray data were processed with simple scaling normalization using Lumi package (http://bioconductor.org/packages/release/bioc/html/lumi.html) [12,13] in Bioconductor [14]. Subsequently, the differentially methylated regions (DMR) between osteosarcoma and normal samples were identified using methyAnalysis with M-value > 1 and false discovery rate (FDR) < 0.01.

M-values were calculated by the formula:

$$M - \text{value} = \log 2 \frac{\text{methylated probe intensity}}{\text{unmethylated probe intensity}} \quad (1)$$

FDR method is also called Benjamini and Hochberg (BH) method [15], used to adjust p value. In detail, the original p values of all genes were ranked in descending order. The maximum p value was assigned as n, and the minimum was assigned as 1. The adjusted p value (FDR) was calculated as followed:

$$\text{FDR} = \text{original p value}^*(n/i) \quad (2)$$

where n represents the number of all genes; i represents the ith p value (from minimum to maximum).

The identified DMR were performed gene annotation to screen the differentially methylated sites located in promoter region. Dkhil et al. [16] have reported that most of the promoters display the changes of DNA methylation in their Ups-regions, which are between +500 and +2000 bp upstream from the transcription start site of genes. Therefore, in consideration of numerous genes in the dataset, the promoter region was defined as 2000 bp upstream of the transcription start site in this study.

2.3. Functional enrichment analysis of differentially methylated genes

Functional enrichment analyses included Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Reactome pathways and Gene Ontology (GO) terms analyses. The differentially methylated genes were performed functional enrichment analysis using GOFunction (http://www.bioconductor.org/packages/release/bioc/html/GOFunction.html) in R package. P value < 0.05 and count (number of significantly enriched genes) ≥ 2 were used as thresholds.

2.4. PPI network construction

PPI analysis can provide new insights into protein function, and uncover the generic organization principles of functional cellular networks [17]. Therefore, we constructed PPI network to further analyze the functions of differentially methylated gene. The Search Tool for the Retrieval of Interacting Genes (STRING) database (http://www.mybiosoftware.com/pathway-analysis/7789) provides the information of both experimental and predicted interactions [18]. With this tool, the differentially methylated gene pairs with significant interactions (combine score > 0.9) were identified and used for construction of PPI network which was visualized using Cytoscape (http://cytoscape.org/) [19]. In the network, the nodes with higher degrees were defined as hub nodes.

2.5. Gene classification analysis

Based on tumor suppressor (TS) genes database [20] and tumor associated genes (TAG) database [21], the known TS genes and oncogenes were selected from the obtained differentially methylated genes.

2.6. TF analysis

On the basis of the TF-target gene pairs in Encode [22,23], the differentially methylated genes corresponding to the differentially methylated TFs were identified. Then according to the obtained TF-target gene interaction pairs, significant TFs (p value < 0.01) were screened using hypergeometric analysis. The formula of hypergeometric distribution was shown as follows:

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} k \in \{0, 1, 2, \ldots m\} \quad (3)$$

where N represents the total number of genes; n represents the number of target genes regulated by TF; M represents the number of differentially methylated genes; k represents the number of differentially methylated genes regulated by TF.

3. Results

3.1. Identification of differentially methylated genes

In the study, a total of 2845 differentially methylated loci were identified in OS samples relative to normal samples, including 1379 hyper-methylation regions and 169 hypo-methylation regions in the promoter region. Additionally, the numbers of hyper- and hypo-methylated loci in enhancers are shown in Table 1, and in exons and other regions are shown in Table 2.

3.2. Functional enrichment analyses of differentially methylated genes

Pathway enrichment analysis revealed that the differentially hyper-methylated genes were significantly enriched in KEGG pathways including Pathways in cancer, Bladder cancer, Neuroactive ligand-receptor interaction, and Peroxisome proliferator-activated receptor (PPAR) signaling pathways (Table 3). The differentially hypo-methylated genes were significantly related to Cytokine-cytokine receptor interaction and Toll-like receptor signaling pathways (Table 4). With regard to Reactome pathways, the differentially hyper-methylated genes were primarily enriched in Gastrin-CREB signaling pathway via PKC and MAPK, and G alpha (q) signaling events pathway, while the differentially hypo-methylated genes were mainly enriched in Defensins, and Alpha-defensins pathways (Table 5). For GO function, the

Table 1
The numbers of hyper- and hypo-methylated loci in enhancer region.

Enhancer	Hyper	Hypo	Total
TRUE	43	163	206
FALSE	1490	1071	2561