Int J Clin Exp Pathol 2020;13(7):1578-1589
www.ijcep.com /ISSN:1936-2625/IJCEP0112561

Original Article

A genome-wide analysis reveals the MeCP2-dependent regulation of genes in BGC-823 cells

Ying-Xun Liu[1,2], Lin-Yuan Li[1], Zhi-Jun Diao[1], Yu-Meng He[1], Ying Chen[3], Ni Hou[2], Ling-Yu Zhao[2], Chen Huang[2]

[1]Key Laboratory of Modern Teaching Technology, Ministry of Education, Shaanxi Normal University, Xi'an 710062, China; [2]Department of Cell Biology and Genetics, School of Basic Medical Sciences, Xi'an Jiaotong University Health Science Center, Xi'an 710061, China; [3]School of Medical Technology, Xuzhou Medical University, Xuzhou 221004, China

Received April 14, 2020; Accepted May 20, 2020; Epub July 1, 2020; Published July 15, 2020

Abstract: Methyl-CpG-binding protein 2 (MeCP2) epigenetically modulates gene expression through genome-wide binding to methylated CpG dinucleotides. This study aimed to evaluate the effect of MeCP2 on the global gene expression profile of human gastric adenocarcinoma to determine the potential molecular mechanism of MeCP2. To identify the gene targets of MeCP2 in gastric cancer cells, we combined the expression microarray and chromatin immunoprecipitation approaches of MeCP2, followed by sequencing (ChIP-seq) to define the MeCP2-binding sites across the whole genome. The methylation levels of the promoters in BGC-823 cells were downloaded from the National Center for Biotechnology Information Gene Expression Omnibus database (GSM1093053). A total of 5,684 ChIP-enriched peaks were identified by comparing IP and Input, using a *p*-value threshold of 10^{-5} in ChIP-seq. The bioinformatics analysis presented a predictive model of the genome-wide MeCP2-binding pattern, in which the MeCP2 binding site is closely related to the transcription start site region in the genome. The results of motif detection showed that the MeCP2-binding regions contained not only the core CpG motif but also the extended poly (A/T) motifs. Finally, an integrative analysis of the sequence features and DNA methylation states revealed that MeCP2's function as a multifunctional transcriptional regulator may not be directly related to the methylation status of the binding site. The first MeCP2 ChIP-seq and gene expression microarray analysis in BGC-823 cells revealed that MeCP2 plays multiple roles in the regulation of gene expression depending on the microenvironment, such as sequence characteristics and the methylation levels of binding sites.

Keywords: Methyl-CpG-binding protein, ChIP-seq, binding site, gastric cancer, si-MeCP2

Introduction

Gastric cancer (GC) is one of the most common human malignant diseases and the second leading cause of cancer mortality worldwide [1]. The reason for the high fatality rate associated with this disease is that most cases of GC are clinically detected only at the advanced stages with distant metastasis. Unfortunately, tumorigenesis and development are complex processes that require the integration of signaling events that occur in distinct locations within a cell. However, little is known about the molecular mechanisms underlying tumor development. Therefore, a highly critical need is to explore the molecular mechanisms and identify the "key" molecular markers related to metastasis in GC, in order to identify new targets for intervention in the metastatic recurrence of GC [2].

DNA methylation is a crucial epigenetic mark with roles in embryogenesis and differentiation, X-inactivation, imprinting, and the repression of viral and repeat sequences [3]. Changes in the patterns of DNA methylation have been implicated in the pathogenesis of several human diseases, including cancer [2, 4]. The DNA methylation of promoter sequences is a repressive epigenetic mark that downregulates gene expression [3].

Methyl-CpG-binding protein 2 (MeCP2), located at Xq28 in humans, is one of the methyl-CpG-binding domain (MBD) proteins, which are important constituents of the DNA methylation

machinery [5, 6]. MeCP2 contains two functional domains: an 85-amino-acid MBD essential for its binding to 5-methylcytosine and a 104-amino-acid transcriptional repression domain that interacts with histone deacetylase. MeCP2 has a number of molecular functions including chromatin organization, RNA splicing, and transcription regulation [7]. MeCP2 has been widely studied in neuronal systems ever since it acquired biomedical importance with the discovery that mutations in its gene determined a profound neurodevelopmental autism spectrum disorder, the Rett syndrome [8-10]. Although MeCP2 is important for neuronal function, several studies suggest that the function of other cell types, particularly gastric cells, is impaired by MeCP2 defects [8]. Recently, a non-neuronal role of MeCP2 has emerged in microRNA processing [11], development [8], apoptosis [12], and tumorigenesis [13, 14]. It is reported that an MeCP2 mutation determines growth disadvantage in hepatocellular carcinoma [14], pancreatic cancer [15], and GC cells [13, 16, 17].

MeCP2 is a methyl-cytosine binding protein that functions as a transcriptional repressor by binding to methylated CpG dinucleotide in promoters. However, recent data suggest that MeCP2 may be a modulator of transcription that can both activate or repress target genes [18, 19]. Early genome-wide expression profiling studies revealed that only a few genes were significantly upregulated in *MECP2*-knockout mice. It was surprising that even some genes appeared to be repressed [20]. Subsequent genome-wide expression analyses in specific brain subregions indicate that MeCP2 primarily activates but also represses transcription. The multiple functions of MeCP2 are closely related to genomic distribution, methylation status of binding sites, and interaction with several partners. Chromatin immunoprecipitation (ChIP)-on-chip assays in the human SH-SY5Y neuronal cell line have revealed that 60% of MeCP2 binding sites lie within the intergenic regions, and many of these are located > 10 kb from the nearest gene [21]. Additionally, an integrated, genome-scale ChIP and expression microarray analysis revealed that 63% of active promoters were bound by MeCP2 and only 6% of these were highly methylated [8, 22]. Generally, these data highlight the role of MeCP2 as a genome modulator, whose functions are cell-type specific.

Moreover, our group reported that MeCP2 expression is significantly upregulated in GC, the expression level was correlated with the clinicopathologic features, and MeCP2 promoted GC cell growth [17, 23]. The role of MeCP2 in regulating the proliferation of prostate cancer and hepatocellular carcinoma cells is strongly supported by evidence. However, little is known about its biological characteristics or molecular mechanisms in human GC. In this study, we described the use of ChIP coupled to next-generation high-throughput sequencing to define MeCP2-binding sites across the whole genome irrespective of whether they coincide with a promoter. Since MeCP2 can regulate the transcription of mRNA and miRNA, we examined the effects of MeCP2 on gene expression in BGC-823 cells treated with MeCP2 small interfering RNA (siRNA). The methylation level of the promoters in BGC-823 cells was reported by Gao et al. using a liquid hybridization capture-based bisulfite sequencing approach, and data were downloaded from the NCBI GEO database (accession number, GSM1093053) [24]. Finally, we performed an integrative analysis of the sequence features and DNA methylation states to reveal the molecular mechanism of MeCP2-mediated epigenetic regulation.

Materials and methods

Cell culture

A stomach adenocarcinoma cell line (BGC-823) was obtained from the GeneChem Corporation (Shanghai, China). The cells were grown at 37°C in a 5% CO_2 incubator, in an RPMI-1640 culture media supplemented with fetal bovine serum (FBS).

ChIP-sequencing

To validate the specific MeCP2 target regions, a MeCP2 ChIP-sequencing (ChIP-seq) analysis was performed on the chromatin prepared from the BGC-823 cells as previously described [9]. Briefly, the BGC-823 cells were grown to a density of 1×10^6/mL in DMEM media supplemented with 10% FBS (Gibco). Subsequently, the cells were treated with 1% formaldehyde for 15 min at room temperature followed by quenching with 125 mM glycine. The cells were washed twice in cold PBS and swollen for 10 min in a hypotonic lysis buffer. The nuclear pellets were collected and lysed in a 1× RIPA buffer. Next,

the nuclear lysates were sonicated using a cell cracker (power 1200 W, 60 cycle for 3×9-s intervals) so that the chromatin was sheared into fragments from 100 to 600 bp. The clarified lysates were divided in half and treated overnight with either anti-MeCP2 (Abcam) rabbit polyclonal antibody or normal rabbit IgG (Santa Cruz, CA). The immunocomplexes were collected with 40 μL protein G agarose (Invitrogen, USA) and slurry for 2 h at 4°C. The immunoprecipitates were successively washed three times with low wash buffer, high wash buffer, and TE buffer. Next, the protein-DNA complexes were eluted from the beads with elution buffer. Subsequently, the elutes were added to RNase A (0.5 mg/mL), Proteinase K (0.2 mg/mL), and NaCl (300 mM) and incubated overnight at 65°C to reverse the cross-linking. The DNA was purified using phenol/chloroform (Invitrogen, USA) extraction and ethanol precipitation and used as the libraries for the RNA sequencing.

The high-throughput sequencing of the MeCP2 ChIP and control libraries was performed using the Illumina HiSeq 2000 sequencing platform with the following steps: 10 ng of DNA libraries were blunt-ended with T4 DNA polymerase and Klenow polymerase. A dA base was added to the 3' end of each strand by Klenow (exo minus) polymerase. Illumina's genomic adapters were ligated to the DNA fragments. PCR amplification was performed to enrich the ligated fragments. An enriched product of approximately 200-400 bp was cut from the gel and purified using a QIAquick Gel Extraction Kit. The completed libraries were quantified using an Agilent 2100 Bioanalyzer. The libraries were denatured with 0.1 M NaOH to generate single-stranded DNA molecules, captured on Illumina flow cells, and amplified in situ. Then, the libraries were sequenced on the Illumina HiSeq 2000 following the TruSeq Rapid SBS Kit protocol.

MeCP2 binding site analysis

After the sequencing platform generated the sequencing images, the stages of image analysis and base calling were performed using Off-Line Basecaller software (OLB V1.8). After passing the Solexa CHASTITY quality filter, the clean reads were aligned to the human genome (UCSC HG19) using Bowtie software (V2.1.0). The aligned reads were used for the peak calling of the ChIP regions using MACS v 1.4.0. The statistically significant ChIP-enriched regions (peaks) were identified by comparing the IP and Input, using a *p*-value threshold of 10^{-5}. The peaks in the samples were annotated using the nearest gene and using the newest UCSC RefSeq database using the ChIPpeakAnno and ChIPseeker package of R/Bioconductor [25, 26]. The signal profile (at 10 bp resolution) with a UCSC WIG file format was generated from the ChIP-seq data, which can be visualized on the UCSC genome browser (http://genome.ucsc.edu/). De novo motif discovery was conducted using the Gibbs motif sampler in CisGenome using a mean motif length of 12 bp and a maximum length of 30 bp. To address the possibility of an alternate MeCP2 binding motif, our unbiased and more precise pan-genomic identification of the > 209 MeCP2 binding sites allows the application of the de novo motif detection tool MEME on the sequences within ± 100 bp of the summits of the MeCP2 peaks with an FDR < 1%.

siRNA transfection

siRNA was predesigned for the MeCP2 gene silencing. Human MeCP2 siRNA (sense 5'-GCUUAAGCAAAGGAAAUCUUU-3', antisense 5'-AGAUUUCCUUUGCUUAAGCUU-3') and negative siRNA (NC-siRNA, sense 5'-UUCUCCGAACGUGUCACGUTT-3', antisense 5'-ACGUGACACGUUCGGAGAATT-3') were chemically synthesized by the Shanghai GenePharma Corporation (Shanghai, China). The BGC-823 cells were cultured for 24 h in plates, and the siRNAs were transiently transfected into the cells using jetPRIME® reagent (Polyplus-transfection) according to the manufacturer's protocol.

Expression microarray hybridization

The total RNA was isolated by homogenizing the cell cultures and processing them using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. Global gene expression analysis was applied using the Human Gene Expression 4×44K v2 Microarray Kit (Agilent Technologies; Cat. No., G4845A). Following the washing steps, the slides were scanned using the Axon GenePix 4000B microarray scanner.

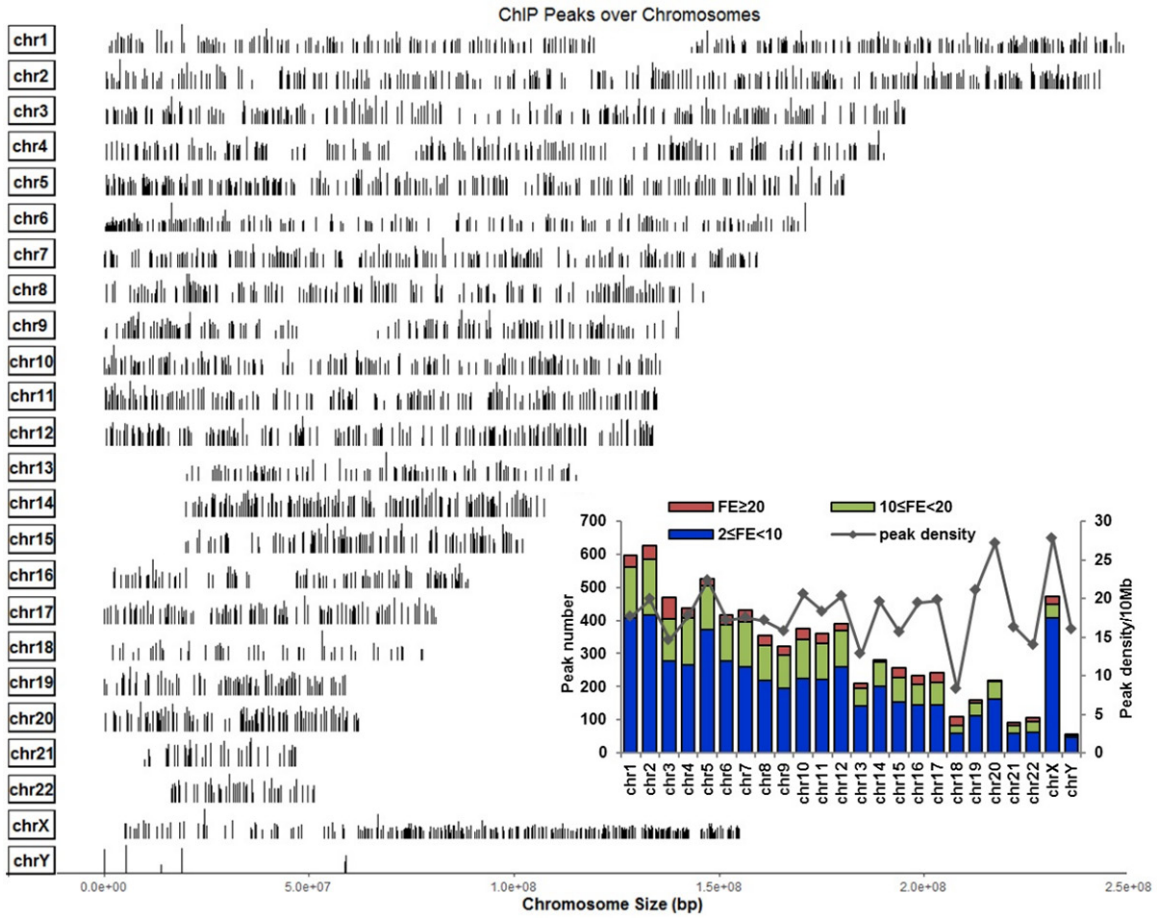

Figure 1. Distribution of the MeCP2 binding sites on the chromosomes. The covplot function calculates the coverage of peak regions over chromosomes using ChIPseeker package of R/Bioconductor. The inset shows the peak numbers and densities on the chromosomes.

Results

MeCP2 peak distribution on chromosomes

Initially, we created a whole-genome map of the MeCP2 binding sites using the ChIP-seq approach in the BGC-823 cells. Sequence reads were generated from Illumina HiSeq 2000, and the image analysis and base calling were performed using Off-Line Basecaller software (OLB V1.8.0). A total of 5,684 ChIP-enriched peaks were identified by comparing MeCP2 IP and Input, using a *p*-value threshold of 10^{-5}.

Generally, our sequencing results obtained 5,089 peaks with a fold enrichment more than twofold. To further clarify the distribution characteristics of MeCP2 peak on the chromosomes, we used the ChIPseeker package in R/Bioconductor to determine the distribution of the MeCP2 peaks on each chromosome. Moreover, the numbers of MeCP2 peaks present on each chromosome were counted. The results showed that MeCP2 peaks were distributed on every chromosome, and the total number of peaks on each chromosome was nearly proportional to the length of the chromosome (**Figure 1**). However, when calculating the peak density on each chromosome, it was found that MeCP2 had the highest binding strength to Chr 20 and Chr X but the lowest binding strength to Chr 13, Chr 18, and Chr 22.

Another feature of MeCP2 distribution on the chromosome is that MeCP2 does not correspond to the centromere of the metacentric chromosome (e.g., Chr 1 and Chr 9) or to the short arm (p) of the acrocentric chromosome (e.g., Chr 13, 14, 15, 21, and 22). These results

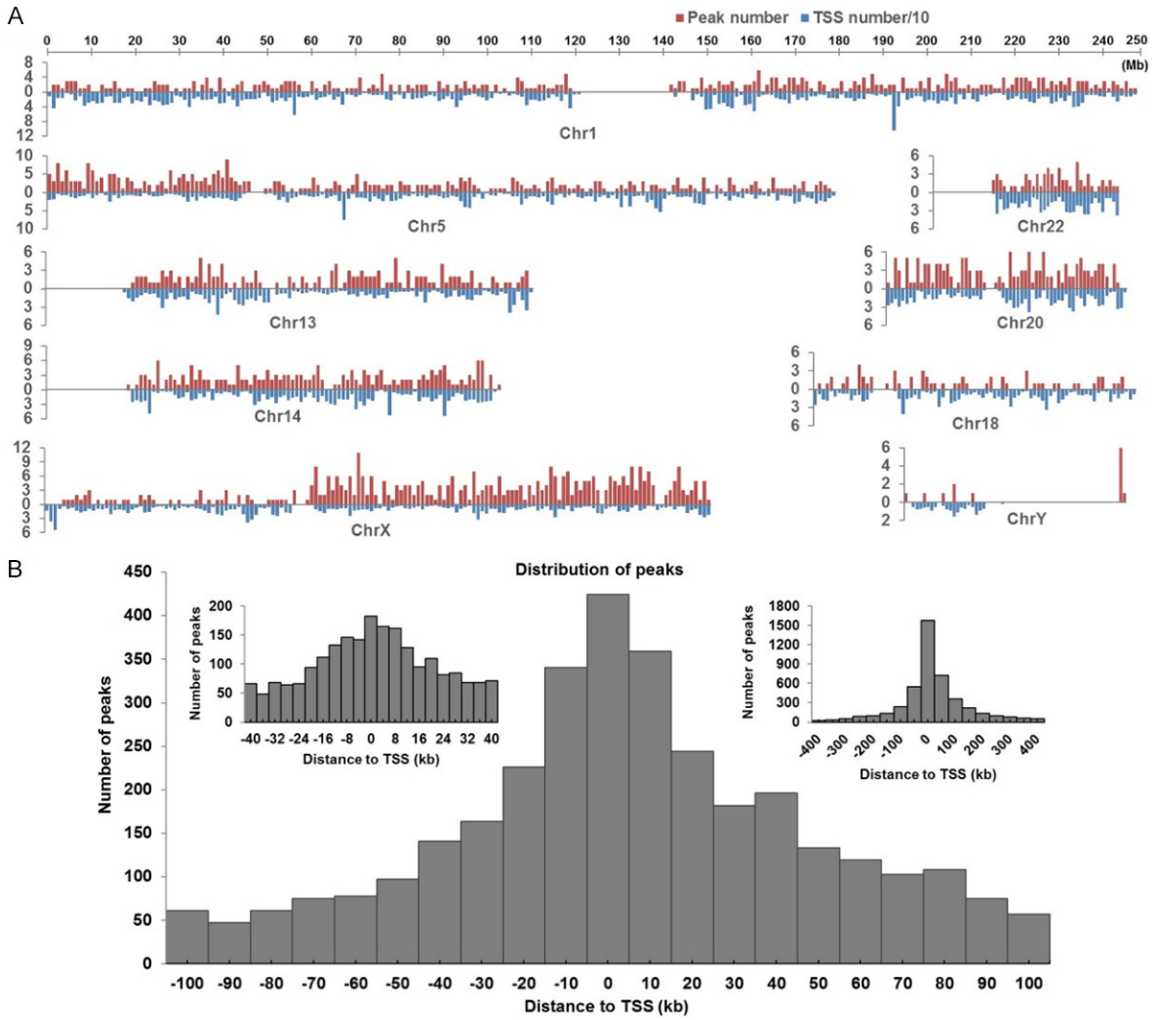

Figure 2. MeCP2 binding is associated with chromosome and TSS. A. The human genome was partitioned into histone modification-poor and aligned at position 0 of the X-axis. The distribution of 4,794 in vivo MeCP2 sites was compared with the distribution of 4,794 randomly selected genomic positions sampled from each chromosome according to the number of TSS. To show the overall distribution trend in limited space, the bars representing the TSS number were shortened by 10 times. B. The histogram was generated from the putative MeCP2 binding regions within 20 kb around the TSS identified in the BGC-823 cells. The plot shows that the MeCP2 binding sites are enriched in the regions more symmetrically around the TSS.

suggest that the MeCP2 binding site is mainly distributed in the euchromatin region, which contains the coding gene. However, the heterochromatin region that does not contain the coding gene, such as the p arm of the acrocentric chromosome, does not contain the MeCP2 binding sites. The results suggest that the MeCP2 binding site is closely related to gene transcription.

Relationship between MeCP2 peak and transcription start site

Since the distribution of MeCP2 is closely related to gene transcription, we further characterized the chromosome distribution of the ChIP peaks around the transcription start site (TSS). Initially, we calculated the number of TSS in the region while calculating the ChIP peak distribution of each chromosome. In **Figure 2A**, the number of MeCP2 peaks and the number of TSS on a chromosome are segmentally counted from the chromosome start to the end in units of 1 Mb. Each chromosome is reduced in proportion to the length of Chr 1 based on the actual length. The red bar graph above the abscissa represents the number of MeCP2 peaks, and the blue bar graph below the abscissa represents the number of TSS in the corre-

sponding area. To equalize the vertical and horizontal coordinates, the number of TSS is expressed as 1/10 of the actual number. The results showed that the MeCP2 peak distribution was closely related to the distribution of TSS on the chromosome. For example, Chr 1 is a metacentric chromosome, and the centromere region is the structural heterochromatin region. This region does not encode the gene and does not contain the peak of MeCP2. The p arm of Chr 22 also belongs to the structural heterochromatin region, which does not contain the peaks of TSS or MeCP2. Similarly, the numbers of TSS and peaks were simultaneously increased in the euchromatin region of the chromosome decreases.

Then, the distribution of the distance of the MeCP2 peaks to the nearest TSS was plotted using the ChIPpeakAnno package. **Figure 2B** shows that the MeCP2 binding sites were more symmetrically enriched in the regions around the TSS, and the peak decreases with a decrease in the distance from upstream and downstream, showing a normal distribution. This may be due to the open state of the chromosomes around the TSS, suggesting that MeCP2 plays an important role in maintaining chromosome morphology.

The relationship between the MeCP2 peak and the annotation genes in whole genome

As mentioned above, the MeCP2 binding site is closely related to the TSS region in the genome. Additionally, the ChIPpeakAnno package was used to summarize the distribution of the peaks over different types of features, such as the exon, intron, enhancer, proximal promoter, 5' UTR, and 3' UTR (**Figure 3**). Initially, we defined the range of -2000 bp upstream to 500 bp downstream of the TSS as the promoter region, and the range of -2 kb to -20 kb upstream of the TSS as the upstream region. Genetic regions beyond the 20 kb upstream of the TSS were considered to be intergenic regions, and the gene body region includes its own exon region, intron region, 3' UTR, and 5' UTR (**Figure 3A**). Please note that one peak might span multiple types of features, leading to the number of annotated features being greater than the total number of input peaks (**Figure 3B**). At the peak centric view, precedence will dictate the annotation order when the peaks span multiple types of features (**Figure 3C**). The statistical results show that the MeCP2 peak is more in the intergenic region, accounting for 53.3% of the whole, followed by the intron region (29.6%). The results indicated that the binding site of MeCP2 is mainly in the non-coding region of the gene, suggesting a large number of genome non-coding regions play a more important role in regulating gene expression. The promoter and upstream regulatory regions occupy 4% and 6%, respectively. Although the number is not dominant and the combination of this region is only a short 20 kb, compared with a large number of non-coding regions, this density is also considerable. This is consistent with the results reported by LaSalle et al. By doing a ChIP-on-chip analysis of a neuroblastoma cell line, they found that more than half of the MeCP2 binding sites are intergenic and only a small number of these reside in the CpG islands [21].

Sequence characteristics of the MeCP2 peaks in the promoter region

To further investigate the sequence specificity of the MeCP2 binding sites on different regions, we performed de novo motif discovery using the DNA FASTA sequences of the MeCP2 peak regions. Since the MeCP2 protein acts as a methyl binding protein, the di-nucleotide methyl-CG (mCG) is the classical DNA recognition sequence. As expected, the most enriched motif identified in this way for the MeCP2 binding regions contained the core CpG motif (**Figure 3D**). From the motif diagram, base C is conserved, and there is a certain change in the next site. It was consistent with the finding that MeCP2 also binds methylated cytosines in the CH context (where H = A/T/C), with a similar affinity both in vitro and in vivo [27, 28]. Interestingly, the MeCP2 peak exhibits the extended poly (A/T) motifs in different binding regions, including promoter, intergenic regions, and gene body. The in vitro selections for the MeCP2-bound DNA-enriched fragments containing A/T bases were adjacent to methyl-CpG. The A/T bases were found to be essential for high-affinity binding at the selected sites and at the known MeCP2 target regions in the Bdnf and Dlx6 genes. These motifs within MeCP2 binding sites may contribute to the transcriptional function in addition to the binding sites [18, 29, 30].

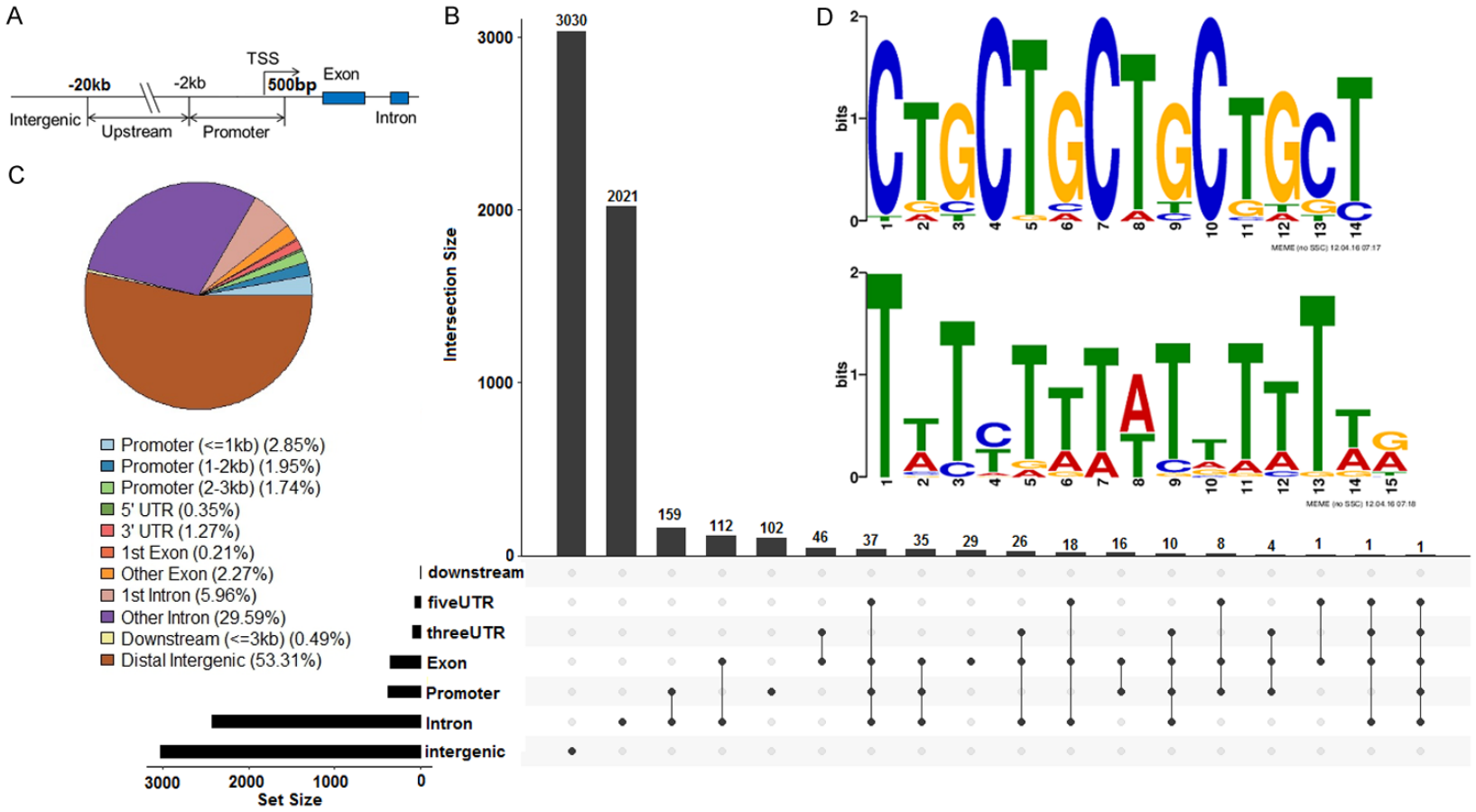

Figure 3. Sequence characteristics of the MeCP2 binding regions. A. The diagram of the gene structure that contains 5 classes based on their distances to UCSC RefSeq genes. B, C. ChIP peak distribution across different genomic regions using the ChIPseeker package of R/Bioconductor. D. Discovery of the novel binding motifs for the MeCP2 binding regions using MEME. The y-axis shows the relative entropy in bits, a measure of the probability that the letter will be found at that position relative to the total information content of the stack, while the x-axis shows the width of the motif in base pairs.

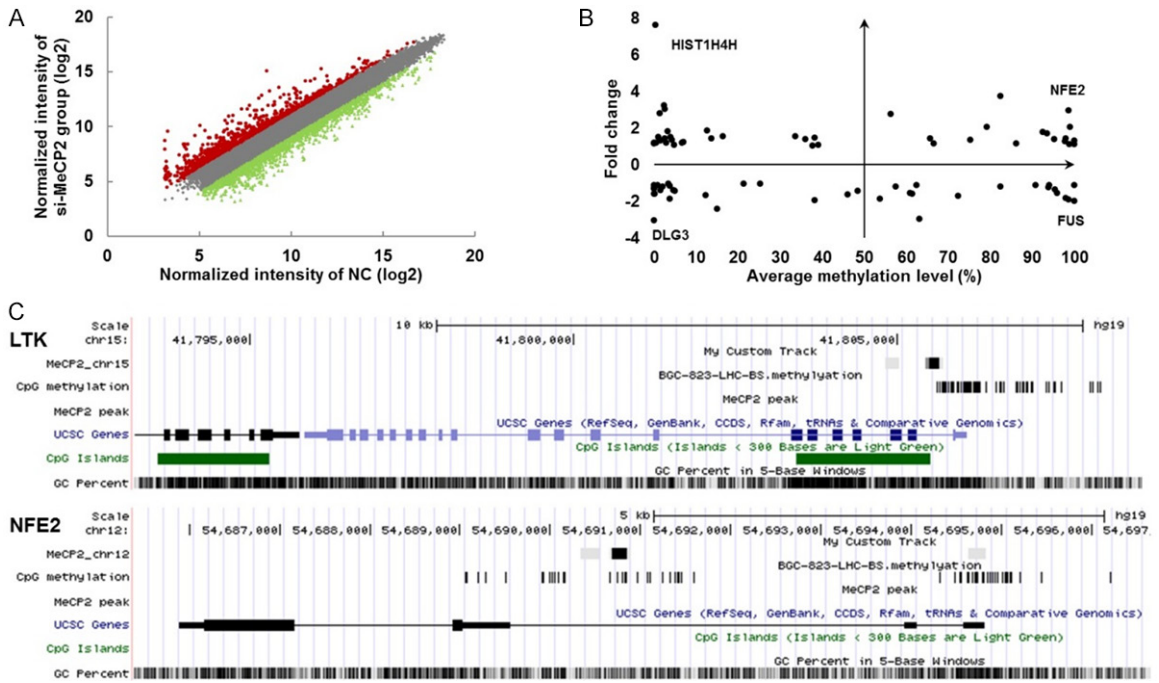

Figure 4. The relationship between the methylation levels of the MeCP2 binding sites within the promoter and gene expression in the BGC-823 cells. A. Scatter plot showing gene expression profile change after MeCP2 siRNA transfection in BGC-823 cells. B. Correlation between the fold change of gene expression and the average methylation level of the promoter region. The methylation data of the promoter region was downloaded from the NCBI GEO database (GSM1093053). C. Examples of MeCP2 peak binding enrichment at specific gene regions. Cumulative DNA sequences co-purified with MeCP2 from BGC-823 cells are depicted as wiggle tracks aligned to the UCSC genome browser (GRCh37/hg19) track of the human genome.

Effect of binding sites within the promoter on gene expression

MeCP2 is a methyl-cytosine binding protein that is proposed to function as a transcriptional repressor by binding to methylated CpG dinucleotides in promoters. However, recent data suggest that MeCP2 may be a modulator of transcription that can both activate or repress target genes. Here, we hypothesized that the MeCP2 occupancy at gene promoters may promote transcription silencing or activation, which is due to the methylation status of the binding sites.

To further explore these possibilities, we analyzed the relationship between the methylation statuses of the gene promoter and gene expression changes after the MeCP2 knockdown in the BGC-823 cells. To investigate the effects of MeCP2 on the gene expression profiles in the BGC-823 cells and to identify the genes affected by MeCP2 in the BGC-823 cells, we depressed the MeCP2 expression using MeCP2 siRNA. After confirmation of the *MeCP2* transcript ablation in the BGC-823 cells, an expression microarray analysis was performed with RNA extracted from gastric cell cultures derived from si-MeCP2 and the control groups. The microarray results demonstrated that 1352 upregulated and 1313 downregulated genes were significantly changed more than twofold, including the MIXL1, THAP11, and HIST1H4H transcripts (**Figure 4A**).

The methylation level of the promoters in BGC-823 cells was reported by Gao et al., and the data were downloaded from the NCBI GEO database (accession number, GSM1093053) [24]. The BED file containing the percentages of methylation at each site in the promoter region was uploaded to the UCSC browser (http://genome-euro.ucsc.edu), and we analyzed the overlap with our 209 peaks located in the promoter region. Then, the methylation level of each CpG site in the promoter region could be visualized on the UCSC genome browser.

Unfortunately, when we analyzed the correlation between methylation levels of promoters

and changes in gene expression, no correlation was observed. Regardless of the methylation level, si-MeCP2 could cause significant up- or downregulation of the gene expression (**Figure 4B**). Additionally, **Figure 4C** showed the detail information about the MeCP2 binding region at the promoter of the LTK and NFE genes. The promoters of these two genes contain high methylation levels, but they have different expression changes.

Discussion

MeCP2 was the first methyl-binding protein and is responsible for several neurological disorders [2, 22]. The molecular details of MeCP2 transcriptional regulation have proven to be more complex than initially assumed, and little is known about how the broad binding pattern of MeCP2 regulates transcription in GC. In the present study, we used next-generation high-throughput sequencing following ChIP to present a genome-wide MeCP2 binding pattern. Furthermore, we describe the impact of MeCP2 knockdown on transcriptional regulation. Finally, our integrative analysis of the sequence features and DNA methylation states revealed that MeCP2 function as a multifunctional transcriptional regulator and may not be directly related to the methylation status of the binding site.

We used BWA tools with the default settings to map these reads to the hg19 human genome reference assembly. Next, the MeCP2 peaks were identified using MACS software. A total of 5,684 ChIP-enriched peaks were identified by comparing the MeCP2 IP and IgG Input. The bioinformatics analysis of the MeCP2-binding genes revealed that MeCP2 has a wide coverage in the human genome. The MeCP2 binding sites are mainly distributed in the euchromatin region, which contains the coding gene (**Figures 1** and **2A**). However, the heterochromatin region that does not contain the coding gene, such as the p arm of the acrocentric chromosome, does not contain MeCP2 binding sites. The bioinformatics results also showed that approximately 53.3% of the MeCP2 binding sites are intergenic, although the binding sites were enriched in regions more symmetrically around the TSS (**Figures 2B** and **3**). These results were consistent with the ChIP-on-chip analysis on a neuroblastoma cell line, which was reported by LaSalle et al. They claimed that more than half of the MeCP2 binding sites are intergenic and that only a small number of them reside in the CpG islands [21]. Moreover, studies have revealed significant differences in the numbers and sizes of the nucleoli and chromocenters in the MeCP2 null mouse neurons compared to those in the WT animals [31]. MeCP2 is able to bind at a genome-wide level, and it has become evident that MeCP2 functions in the chromatin architecture and genome organization [18, 31].

MeCP2 has a high affinity for methylated cytosines that are followed by a guanine nucleotide (mCG) [5, 32]. To address the possibility of an alternate MeCP2 binding motif, we performed a de novo motif discovery on the MeCP2-binding sequences. Our ChIP-seq data suggest that the most enriched motif identified in this way for MeCP2 binding regions contained not only the core CpG motif but also the extended poly (A/T) motifs (**Figure 3D**). Recently, it was found that MeCP2 also binds methylated cytosines in the CH context (where H = A/T/C), with a similar affinity both in vitro and in vivo [27, 28, 32]. Thus, the binding of MeCP2 at mCG, mCH, and hmC along with the accompanying alteration in chromatin architecture can collectively play a role in regulating gene expression in mature neurons [19]. Interestingly, the MeCP2 peak exhibits the extended poly (A/T) motifs in different binding regions, including promoter, intergenic regions, and gene body. These motifs may contribute to a transcriptional function in addition to the binding sites (20, 29, 30). An artificial selection of the MeCP2 binding sites in vitro using methyl-SELEX demonstrated that MeCP2 requires an A/T run of ≥ 4 base pairs adjacent to the methyl-CpG for efficient DNA binding. We propose that the presence of an A/T run is essential at the biologically relevant MeCP2 binding sites and may help predict MeCP2 interaction sites [29, 30].

Generally, methylation of the promoter region inhibits the transcription of several mRNAs and miRNAs, which play an important role in tumor development. Previous studies have reported that the genetic disruption of DNA methylation in cancer cell lines induces the upregulation of substantial numbers of mRNAs and miRNAs [2, 13, 33]. The published results from our group demonstrated that 5'-Aza-dC treatment can induce demethylation in the promoter region and restore gene expression [34]. Unfortunately, the apparent dichotomy of MeCP2 functions

(methylation dependence vs independence) is highly controversial. Previous studies suggested that MeCP2 is a transcriptional repressor targeted to specific genes via DNA methylation [7, 18]. However, subsequent studies have proposed that MeCP2 acts as a transcriptional activator by recruiting the transcription factor CREB to specific genes. Furthermore, there have been reports that MeCP2 binding to DNA may not depend on methylation, as studies using in vitro chromatin assembly suggested that MeCP2 can bind to both methylated and nonmethylated DNAs and mediate nucleosomal compaction [5]. To distinguish these alternative hypotheses for MeCP2 function, we analyzed the relationship between the methylation status of the gene promoter and gene expression changes after the MeCP2 knockdown in the BGC-823 cells. We utilized the NCBI GEO database to determine the methylation levels of the MeCP2-binding promoter region. The results demonstrated that MeCP2 does not selectively bind methylated promoters but is bound at many unmethylated promoters and has a similar regulation pattern to gene expression (**Figure 4C**). Consistent with this methylation-independent binding manner, Skene performed a ChIP analysis and found MeCP2 to be globally distributed and to track the density of meCpG in the neuron genome. Bisulfite sequencing of the immunoprecipitated DNA confirmed highly selective binding to the methylated DNA sequences. The author claimed that it was the key neuron-specific role of MeCP2, which was not observed in the glia or other tissues [35].

Therefore, using next-generation RNA sequencing and gene expression microarray, the present study has successfully identified a unique set of genes responsive to MeCP2 in BGC-823 cells. The bioinformatics analysis of the ChIP-enriched peaks presents a high-resolution genome-wide mapping of the MeCP2 binding sites. Our integrative analysis of the sequence features and DNA methylation states revealed that MeCP2's function as a multifunctional transcriptional regulator may not be directly related to the methylation status of the binding site. Generally, depending on the microenvironment, MeCP2 can function as a transcript repressor and also act as a gene expression activator. These findings also explore the complexity of MeCP2-mediated epigenetic regulation in GC.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (31400730) and the Fundamental Research Funds for the Central Universities (GK201903102).

Disclosure of conflict of interest

None.

Address correspondence to: Chen Huang and Ling-Yu Zhao, Department of Cell Biology and Genetics, School of Basic Medical Sciences, Xi'an Jiaotong University Health Science Center, Xi'an 710061, China. Tel: +86-29-82657723; Fax: +86-29-8265-7723; E-mail: hchen@xjtu.edu.cn (CH); Tel: +86-29-82655190; E-mail: zhaolingyu@xjtu.edu.cn (LYZ)

References

[1] Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D. Global cancer statistics. CA Cancer J Clin 2011; 61: 69-90.

[2] Suzuki H, Maruyama R, Yamamoto E and Kai M. DNA methylation and microRNA dysregulation in cancer. Mol Oncol 2012; 6: 567-578.

[3] Jjingo D, Conley AB, Yi SV, Lunyak VV and Jordan IK. On the presence and role of human gene-body DNA methylation. Oncotarget 2012; 3: 462-474.

[4] Yang X, Han H, De Carvalho DD, Lay FD, Jones PA and Liang G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell 2014; 26: 577-590.

[5] Chen L, Chen K, Lavery LA, Baker SA, Shaw CA, Li W and Zoghbi HY. MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. Proc Natl Acad Sci U S A 2015; 112: 5509-5514.

[6] Hansen JC, Wexler BB, Rogers DJ, Hite KC, Panchenko T, Ajith S and Black BE. DNA binding restricts the intrinsic conformational flexibility of methyl CpG binding protein 2 (MeCP2). J Biol Chem 2011; 286: 18938-18948.

[7] Chen WG, Chang Q, Lin Y, Meissner A, West AE, Griffith EC, Jaenisch R and Greenberg ME. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. Science 2003; 302: 885-889.

[8] Cohen S, Gabel HW, Hemberg M, Hutchinson AN, Sadacca LA, Ebert DH, Harmin DA, Greenberg RS, Verdine VK, Zhou Z, Wetsel WC, West AE and Greenberg ME. Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function. Neuron 2011; 72: 72-85.

[9] Yasui DH, Xu H, Dunaway KW, LaSalle JM, Jin LW and Maezawa I. MeCP2 modulates gene expression pathways in astrocytes. Mol Autism 2013; 4: 3.
[10] Urdinguio RG, Fernandez AF, Lopez-Nieva P, Rossi S, Huertas D, Kulis M, Liu CG, Croce C, Calin GA and Esteller M. Disrupted microRNA expression caused by Mecp2 loss in a mouse model of Rett syndrome. Epigenetics 2010; 5: 656-663.
[11] Cheng TL, Wang Z, Liao Q, Zhu Y, Zhou WH, Xu W and Qiu Z. MeCP2 suppresses nuclear microRNA processing and dendritic growth by regulating the DGCR8/Drosha complex. Dev Cell 2014; 28: 547-560.
[12] Bracaglia G, Conca B, Bergo A, Rusconi L, Zhou Z, Greenberg ME, Landsberger N, Soddu S and Kilstrup-Nielsen C. Methyl-CpG-binding protein 2 is phosphorylated by homeodomain-interacting protein kinase 2 and contributes to apoptosis. EMBO Rep 2009; 10: 1327-1333.
[13] Tong D, Zhao L, He K, Sun H, Cai D, Ni L, Sun R, Chang S, Song T and Huang C. MECP2 promotes the growth of gastric cancer cells by suppressing miR-338-mediated antiproliferative effect. Oncotarget 2016; 7: 34845-34859.
[14] Zhao LY, Zhang J, Guo B, Yang J, Han J, Zhao XG, Wang XF, Liu LY, Li ZF, Song TS and Huang C. MECP2 promotes cell proliferation by activating ERK1/2 and inhibiting p38 activity in human hepatocellular carcinoma HEPG2 cells. Cell Mol Biol (Noisy-le-grand) 2013; Suppl 59: OL1876-1881.
[15] Xu M, Bian S, Li J, He J, Chen H, Ge L, Jiao Z, Zhang Y, Peng W, Du F, Mo Y and Gong A. MeCP2 suppresses LIN28A expression via binding to its methylated-CpG islands in pancreatic cancer cells. Oncotarget 2016; 7: 14476-14485.
[16] Wada R, Akiyama Y, Hashimoto Y, Fukamachi H and Yuasa Y. miR-212 is downregulated and suppresses methyl-CpG-binding protein MeCP2 in human gastric cancer. Int J Cancer 2010; 127: 1106-1114.
[17] Zhao LY, Tong DD, Xue M, Ma HL, Liu SY, Yang J, Liu YX, Guo B, Ni L, Liu LY, Qin YN, Wang LM, Zhao XG and Huang C. MeCP2, a target of miR-638, facilitates gastric cancer cell proliferation through activation of the MEK1/2-ERK1/2 signaling pathway by upregulating GIT1. Oncogenesis 2017; 6: e368.
[18] Della Ragione F, Filosa S, Scalabrì F and D'Esposito M. MeCP2 as a genome-wide modulator: the renewal of an old story. Front Genet 2012; 3: 181.
[19] Della Ragione F, Vacca M, Fioriniello S, Pepe G and D'Esposito M. MECP2, a multi-talented modulator of chromatin architecture. Brief Funct Genomics 2016; 15: 420-431.
[20] Vacca M, Tripathi KP, Speranza L, Aiese Cigliano R, Scalabrì F, Marracino F, Madonna M, Sanseverino W, Perrone-Capano C, Guarracino MR and D'Esposito M. Effects of Mecp2 loss of function in embryonic cortical neurons: a bioinformatics strategy to sort out non-neuronal cells variability from transcriptome profiling. BMC Bioinformatics 2016; 17 Suppl 2: 14.
[21] Yasui DH, Peddada S, Bieda MC, Vallero RO, Hogart A, Nagarajan RP, Thatcher KN, Farnham PJ and LaSalle JM. Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. Proc Natl Acad Sci U S A 2007; 104: 19416-21.
[22] Rube HT, Lee W, Hejna M, Chen H, Yasui DH, Hess JF, Lasalle JM, Song JS and Gong Q. Sequence features accurately predict genome-wide MeCP2 binding in vivo. Nat Commun 2016; 7: 11025.
[23] Zhao L, Liu Y, Tong D, Qin Y, Yang J, Xue M, Du N, Liu L, Guo B, Hou N, Han J, Liu S, Liu N, Zhao X, Wang L, Chen Y and Huang C. MeCP2 promotes gastric cancer progression through regulating FOXF1/Wnt5a/β-Catenin and MYOD1/Caspase-3 signaling pathways. EBioMedicine 2017; 16: 87-100.
[24] Gao F, Wang J, Ji G, Liu S, Yao Y, Wang T, Wu H, Xia Y, Gong D, Jiang H, Yang H and Zhang X. Clustering of cancer cell lines using a promoter-targeted liquid hybridization capture-based bisulfite sequencing approach. Technol Cancer Res Treat 2015; 14: 383-394.
[25] Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS and Green MR. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics 2010; 11: 237.
[26] Yu G, Wang LG and He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 2015; 31: 2382-2383.
[27] Lagger S, Connelly JC, Schweikert G, Webb S, Selfridge J, Ramsahoye BH, Yu M, He C, Sanguinetti G, Sowers LC, Walkinshaw MD and Bird A. MeCP2 recognizes cytosine methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain. PLoS Genet 2017; 13: e1006793.
[28] Liu K, Xu C, Lei M, Yang A, Loppnau P, Hughes TR and Min J. Structural basis for the ability of MBD domains to bind methyl-CG and TG sites in DNA. J Biol Chem 2018; 293: 7344-7354.
[29] Lyst MJ, Connelly J, Merusi C and Bird A. Sequence-specific DNA binding by AT-hook motifs in MeCP2. FEBS Lett 2016; 590: 2927-2933.
[30] Klose RJ, Sarraf SA, Schmiedeberg L, McDermott SM, Stancheva I and Bird AP. DNA binding selectivity of MeCP2 due to a requirement for

A/T sequences adjacent to Methyl-CpG. Mol Cell 2005; 19: 667-678.
[31] Singleton MK, Gonzales ML, Leung KN, Yasui DH, Schroeder DI, Dunaway K and LaSalle JM. MeCP2 is required for global heterochromatic and nucleolar changes during activity-dependent neuronal maturation. Neurobiol Dis 2011; 43: 190-200.
[32] Mellen M, Ayata P, Dewell S, Kriaucionis S and Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell 2012; 151: 1417-1430.
[33] Guo B, Zhao Z, Wang Z, Li Q, Wang X, Wang W, Song T and Huang C. MicroRNA-302b-3p suppresses cell proliferation through AKT pathway by targeting IGF-1R in human gastric cancer. Cell Physiol Biochem 2017; 42: 1701-1711.
[34] Guo B, Zhang J, Li Q, Zhao Z, Wang W, Zhou K, Wang X, Tong D, Zhao L, Yang J and Huang C. Hypermethylation of miR-338-3p and impact of its suppression on cell metastasis through N-cadherin accumulation at the cell -cell junction and degradation of MMP in gastric cancer. Cell Physiol Biochem 2018; 50: 411-425.
[35] Skene PJ, Illingworth RS, Webb S, Kerr AR, James KD, Turner DJ, Andrews R and Bird AP. Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Mol Cell 2010; 37: 457-468.