Abstract. Osteosarcoma is a common bone tumor which is affected by E2, the most representative estrogen. Gene regulation function of E2 is highly dependent on estrogen receptor. The purpose of this study was to explore the gene regulation patterns of E2 through estrogen receptor \( \alpha \) (ESR1) in osteosarcoma based on the combined analysis of ChIP-seq and gene microarray. All of the datasets were downloaded from the Gene Expression Omnibus (GEO). Differential expression genes (DEGs) in E2 treated U2OS cells expressing ESR1 (U2OS-ER\( \alpha \)) compared with those treated with vehicle were obtained based on R programming software. ESR1-specific binding sites (peaks) in E2 treated U2OS cells were identified through MACS. Overlaps between DEGs and ESR1 target genes which contained peaks in promoters were considered as reliable E2-mediated genes through ESR1 in osteosarcoma. Moreover, we conducted miRNA-Gene regulation analysis for those genes through miRWalk database to identify potential therapeutic targets for the genes. Functional enrichment analysis of DEGs indicated their potential involvement in cancer, and cell activity-related processes. Fifteen overlaps were identified between DEGs and target genes of ESR1, of which 12 were found to be regulated by miRNA. Several known estrogen response genes and novel genes were obtained in this study and they might provide potential therapeutic targets for osteosarcoma.

Introduction

ESR1, also known as ER-alpha or ER\( \alpha \), is an important estrogen receptor (ER) and involves in the gene regulation in various diseases and biological processes, including breast cancer (1,2), osteosarcoma (3) and cell growth (4). 17\( \beta \)-estradiol (E2) is one of the most representative estrogens responsible for the development and reproductive capability of female characteristics (5). Besides, it participates in the progression of many diseases, for example, E2 could regulate the proliferation of breast cancer cells through focal adhesion and chemokine signaling pathways (6); through autophagy and apoptosis pathways, E2 plays important roles in the regulation of renal cell carcinoma growth (7). Furthermore, ER, including ER\( \alpha \) and ER\( \beta \), are important media through which E2 involves the regulation of many biological processes. So, exploration of interactions between ER and E2 would promote our understanding of many types of complex diseases and the identification of their potential therapeutic targets.

Osteosarcoma is one of the most common malignant neoplasms in children and adolescents with ~20% 5-year overall survival rate after chemotherapy (8-10). In primary osteosarcoma, adequate surgical margins could significantly improve the prognosis, whereas, it is still a challenging surgical technique (11). Moreover, metastasis greatly affects prognosis for osteosarcoma due to the propensity of early metastatic trend and ~80% of osteosarcoma patients have metastasis at diagnosis (12,13). Understanding its mechanisms and screening of potential biomarkers would be helpful for its early diagnosis and improvement of prognosis. The development of biotechnologies have promoted the identification of diagnosis and therapeutic targets for osteosarcoma. For example, microRNA-491-5p was considered to suppress cell proliferation and promote apoptosis via FOXP4 in the report of Yin et al (14) through the quantification of miR-491-5p expression and cell apoptosis; through immunohistochemistry and statistic analysis, Zhao et al (15) identified HIF1 as an independent prognostic biomarker in osteosarcoma; furthermore, Zhou et al (16) even indicated that the level of miR-199a-5p in serum/plasma might serve as biomarker for osteosarcoma.
Estrone and estriol, also known as E1 and E3, are another type of estrogens which have been proved to be involved in the progression of several diseases, including osteosarcoma. For example, Mochizuki et al proved that E3 is compatible with E2 to induce the expression of IGF-1 mRNA and then affect the proliferative potential of osteosarcoma cells (17). Besides, E1 sulfate represents the most abundant estrogen in circulation of adults and plays a role in human bone maturation and homeostasis (18), and it has been shown to influence the progression of many types of cancers, such as colorectal cancer (19) and breast cancer (20), so it might be also associated with the development of osteosarcoma.

There are also several studies on E2 and ESR1 in osteosarcoma (3,21,22). While, no study has focused on the genome binding profiles of ESR1 with the presence of E2 and the E2-regulated genes through ESR1 simultaneously. In this study, a combined analysis of ESR1 chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) with the presence of E2 and gene expression profiles of E2-regulated through ESR1 was conducted. This would induce several potential target genes of ESR1 in osteosarcoma, and here, the E2-regulated genes in balance. This would be helpful for the identification of E2 targeted molecules in osteosarcoma and the development of corresponding drugs.

**Materials and methods**

**Genome binding and expression profile datasets.** All of the datasets in this study were obtained from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). For the analysis of genome binding of ESR1, the data of Chen et al were used with the accession number of GSE26110 which contains the SRC-1 and ESR1 genome-wide binding profiles with or without the presence of E2 in U2OS osteosarcoma cells. Another dataset (GSE1153), deposited by Stossi et al (22) was composed of U2OS cells expressing ERα (ESR1) or ERβ and treated with vehicle or the same amount of E2 for 4, 8, 24 and 48 h, was used for the determination of E2 targeted genes. The transcriptome expression level of the samples were quantified via GPL91 [HG_U95a] Affymetrix Human Genome U95A Array, a commercial affymetrix gene microarray.

**ChIP-seq data analysis.** We first conducted quality control for the raw sequencing data in fastq with FastQC, a quality control tool for high-throughput sequence data. Bowtie 2 (23) was used for the fast and sensitive read alignment with the default parameters setting after the removal of reads with poor quality, i.e., length <20 and more than 3 bases with the quality score <20 in this study. After the removal of duplicate alignments, the binding sites, referred as peak hereafter, of ESR1 were identified via the Model-based analysis for ChIP-seq (MACS) version 2 (24), a widely used peak caller for transcription factor and histone modifier developed by Feng et al. Finally, ChIPseeker (25) and deepTools (26) were used for the assignment of genomic features, such as gene ID and relative location to transcription start site (TSS) to the peaks and visualization of binding profiles of ESR1 respectively.
Microarray data analysis. Gene expression profile analysis mainly involved two steps, i.e., preprocessing and identification of differential expression genes (DEGs). Briefly, raw CEL data was imported to R software and normalized through affy package (27) and the expression value was logarithmically transformed based on 2 and summarized to gene level; the identification of DEGs was performed through limma package (28) with the thresholds of fold change >1.41 and Bonferroni adjusted P-value <0.05.

Functional enrichment analysis. To explore the functions involving DEGs, we performed functional enrichment analysis through the Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) (29). We considered Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways with P-value <0.05 and minimum gene hits >2 as significantly enriched.

miRNA-Gene network analysis. In this study, the overlaps between target genes of ESR1 in U2OS cells with the presence of E2 and DEGs in E2 treated U2OS cells with expressed ESR1 were considered E2-regulated genes through ESR1 in osteosarcoma. To further explore their roles in osteosarcoma, we performed miRNA-Gene regulation analysis through miRWalk (http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/) (30) and the regulation relationships were screened out only when they were found in four out of the miRanda, miRDB, miRWalk, RNA22 and TargetScan databases. Besides, Cytoscape (31), an open source platform for complex network analysis and visualization, was used for the presentation of miRNA-Gene network.

Results

ChIP-seq data analysis. Quality control indicated high quality reads suitable for the following analysis (Fig. 1A).
Read alignments resulted in 88.76% (11,409,959 out of 12,854,413 reads) and 75.98% (9,744,663 out of 12,825,873 reads) overall alignment rate for the ChIP-seq data of ESR1 and IgG control, respectively. Through MACS, we obtained a total of 45,355 ESR1-specific binding peaks and their genomic distribution are shown in Fig. 1B. Moreover, 7,712 peaks were assigned to the promoters of 4,541 unique genes, i.e., up- and down-stream 3,000 bp surrounding the TSS, which were considered as the most confident target genes of ESR1 and abbreviated as PGenes in this study. Besides, significant enrichment of ESR1 binding peaks was obtained surrounding TSS of its target genes and shown in Fig. 1C.

Microarray data analysis. Preprocessing step resulted in comparable overall expression values among all of the samples (Fig. 2A). DEGs in 4 h (DEG4), 8 h (DEG8), 24 h (DEG24) and 48 h (DEG48) E2 treated U2OS cells expressing ESR1 (U2OS-ERα) compared with U2OS-ERα treated with vehicle were obtained and cross analyzed. As a result, we identified 216, 240, 654 and 649 hits for DEG4, DEG8, DEG24 and DEG48, respectively, and 34 overlaps were also obtained (Fig. 2B). The 34 overlaps are shown in Table I. Fig. 2C illustrated the supervised two-way hierarchical clustering of the 34 overlapped genes and samples. As expected, 24 and 48 h E2 treated U2OS-ERα samples were clustered together and the 4 h, and 8 h treated, and vehicle treated U2OS-ERα were separately clustered in their own groups.

Enriched functions. We identified 50, 44, 175 and 171 significantly enriched GO terms for DEG4, DEG8, DEG24 and
DEG48, respectively. Most of the GO terms were closely related with the processes of cancer progression, cell adhesion, RNA binding and so on. Besides, several estrogen, cancer and inflammatory associated KEGG pathways were also obtained and shown in Fig. 3.

**miRNA-Gene network.** Cross analysis of PGenes and the 34 overlapped DEGs identified 15 overlaps, i.e., RET, TRPM1, CA12, KRT19, GREB1, FL13A1, TACC1, LRIG1, ADCY8, TMOD1, SLC30A1, NLRP3, WNK1, ITGA6 and ENPP2. Three examples of ESR1 binding peaks in the promoter of
miRNA is a kind of non-coding RNA with the length of ~22 nt. It can regulate gene expression through preventing transcription or degradating protein. We further explored the miRNA-Gene regulation relationships of the 15 overlaps for the comprehensive understanding of gene regulation patterns in E2 treated U2OS-ER\(\alpha\) cells. As a result, 12 out of the 15 overlaps except CA12, KRT19 and ADCY8 were found to be regulated by one or more miRNAs. Fig. 5 illustrates the miRNA-Gene regulation network.

Discussion

Osteosarcoma is the most common bone tumor with high morbidity and mortality. E2 was found to play important roles in osteosarcoma through regulating physiological processes and gene expression via ESR1 by many studies (3,32). While, the combined analysis of E2-mediated gene expression and regulation is still limited. In this study, we conducted gene microarray and ChIP-seq analysis in E2 treated U2OS-ER\(\alpha\) cell lines, which would be useful to promote our understanding of the E2-mediated gene regulation in osteosarcoma.

E2-activated ESR1 could directly or indirectly bind to its target genes and then regulate the particular biological processes or gene expression. In osteosarcoma, E2 is necessary for the translocation of ESR1 to the nuclei. For example, Hu et al showed that ESR1 could promote the binding of Sp1 to the promoter of MALAT1, a long non-coding RNA, in U2OS cells to affect physiological processes only with E2 stimulation (32); Tübel et al even proved that the methylation status of ESR1 was affected by E2 (3). In this study, the ChIP-seq analysis of E2 stimulated U2OS-ER\(\alpha\) cell identified the potential target genes of E2 mediated through ESR1. Besides, cross analysis with gene microarray could further screen out more reliable target genes for osteosarcoma.

Functional enrichment analysis resulted in the identification of cancer, cell activity-related KEGG pathways, which further confirmed our results. Estrogen signaling pathway was
found to be significantly enriched in DEG4 and DEG48, which indicated roles of E2 in the activation of ESR1.

E2 could also regulate the expression level of microRNA (miRNA) and affect cell proliferation, migration and invasion in osteosarcoma. In turn, E2-mediated genes might also be regulated by miRNA. In this study, we performed miRNA- Gene regulation analysis for overlaps between PGenes and overlapped genes among the four lists of DEGs. Finally, 12 out of the 15 overlaps were found to be regulated by miRNA and ACC1 with the highest degree. ACC1 could act as anti- and pro-apoptotic regulators which involves in several cellular activities, including embryonic development, homeostasis, and tumorigenesis. Besides, it has been proved as potential biomarker in many types of cancers, such as gliomas, and melanoma (33, 34). Significant enrichment of ESR1-specific binding peaks were identified in the promoter of WNK1 (Fig. 4, bottom), which was regulated by 31 miRNAs in this study. WNK1 and ESR1 were found to have differential expression in osteoporosis simultaneously (35). GREB1 was found with differential expression in U2OS-ERα cells at 4, 8, 24 and 48 h after E2 treated and regulated by E2 through ESR1. Consistent with our study, Lin and Lei reported that GREB1 was a estrogen-responsive gene and with two canonical estrogen response elements (36).

In conclusion, this study identified the potential E2-mediated genes through ESR1 in osteosarcoma via the combined analysis of ChIP-seq and gene microarray analysis. It could be helpful for the identification of therapeutic targets for osteosarcoma.

References

1. Fodorova D, Toporova L, Macejova D, Lastovickova M, Brtko J: Melanoma (33, 34). Significant enrichment of ESR1-specific biomarker in many types of cancers, such as gliomas, and tumorigenesis. Besides, it has been proved as potential biomarker in many types of cancers, such as gliomas, and melanoma.


