BMC Cancer

Construction of an integrated human osteosarcoma database, HOsDb, based on literature mining, microarray analysis, and database retrieval

Yifu Sun[1†], Lishan Wang[2†], Changkuan Li[1], Rui Gu[1], Weidong Zang[2], Wei Song[2] and Peng Xia[3*]

Abstract

Background: Osteosarcoma (OS) is the most frequent primary malignancy of bone with a high incidence in adolescence. This study aimed to construct a publicly available, integrated database of human OS, named HOsDb.

Methods: Microarray data, current databases, and a literature search of PubMed were used to extract information relevant to human OS-related genes and their transcription factors (TFs) and single nucleotide polymorphisms (SNPs), as well as methylation sites and microRNAs (miRNAs). This information was collated for constructing the HOsDb.

Results: In total, we identified 7191 OS tumor-related genes, 763 OS metastasis-related genes, and 1589 OS drug-related genes, corresponding to 190,362, 21,131, and 41,135 gene-TF pairs, respectively, 3,749,490, 358,361, and 767,674 gene-miRNA pairs, respectively; and 28,386, 2532, and 3943 SNPs, respectively. Additionally, 240 OS-related miRNAs, 1695 genes with copy number variations in OS, and 18 genes with methylation sites in OS were identified. These data were collated to construct the HOsDb, which is available at www.hosdatabase.com. Users can search OS-related molecules using this database.

Conclusion: The HOsDb provides a platform that is comprehensive, quick, and easily accessible, and it will enrich our current knowledge of OS.

Keywords: Osteosarcoma, HOsDb, www.hosdatabase.com

Background

Osteosarcoma (OS), the most frequent primary malignancy of bone, commonly occurs in the metaphyseal region of the long bones, developing at sites of rapid bone growth [1]. OS commonly affects children, adolescents, and young adults. The annual incidence of OS in the general population is 2–3/million/year, while in adolescence, especially from 15 to 19 years of age, OS incidence reaches 8–11/million/year [2]. OS accounts for 15% of all solid extracranial cancers in people aged 15–19 years [3]. OS can be divided into several subtypes, such as osteoblastic, chondroblastic, fibroblastic, small cell, telangiectatic, high-grade surface, extra-skeletal, and other lower-grade forms, including periosteal and parosteal [4]. Some OS cases are likely to have a genetic basis, and numerous hereditary disorders associated with germline alterations of tumor suppressor genes have been found in patients with OS, such as hereditary retinoblastoma [5] and Li-Fraumeni cancer family syndrome [6, 7]. However, the mechanisms underlying the pathogenesis of OS remain largely unclear.

* Correspondence: xuanliuuil@163.com
†Yifu Sun and Lishan Wang are the co-first authors
³Department of Orthopedics, The Second Hospital of Jilin University, No.218 Ziqiang Street, Changchun 130022, China
Full list of author information is available at the end of the article

Sun *et al. BMC Cancer* (2020) 20:390

Page 2 of 9

Many databases have been developed to investigate the association between certain molecules of interest and disease pathogenesis from different perspectives. For instance, Online Mendelian Inheritance in Man (OMIM) [8] contains information on the relationship between the phenotype and genotype of all known Mendelian disorders. Wikigenes [9] is a portal that provides information about genes, proteins, chemical compounds and their reported associations with various diseases. The miR2Disease [10] and Human microRNA Disease Database (HMDD) [11] aim to provide comprehensive collection of microRNAs (miRNAs) associated with various human diseases. MethyCancer [12] contains highly integrated data regarding cancer-related genes, DNA methylation sites, and information on cancer from public resources. TRANSFAC is a database of transcription factors (TFs), which offers an integrated system for predicting gene expression regulation [13]. Although research data regarding OS has accumulated during the past decades, to the best of our knowledge, there is only one available database specifically focusing on OS molecular biology, called Osteosarcoma Database [14]. Nevertheless, only 911 OS-associated genes and 81 miRNAs collected through manual literature mining are included in this database, and there is no information available regarding other OS-related molecules, such as TFs or methylation sites [14]. The development of high-throughput laboratory techniques, such as microarray analysis, has enabled generation of large quantities of data associated with OS,

which are an important resource for exploration of potential OS-related molecules, including genes, miRNAs, and copy number variations (CNVs) [15–18]. While these data provide insight into certain aspects of OS, they are not assembled together in a structured format. Thus, there is a need to establish an integrated, OS-specific database or platform of OS-related genes, TFs, methylation sites, and miRNAs.

We collected detailed OS-related data, including OS-related genes, TFs, single nucleotide polymorphisms (SNPs), miRNAs, methylation sites, and CNVs by analyzing several microarray deposits in the Gene Expression Omnibus (GEO) data repository, searching current databases, and mining the literature in PubMed. Using these data, we aimed to construct a publicly available, integrated database of human OS to facilitate the exploration of human OS-related molecules and create a unique resource for research into this disease.

Construction and content
Database construction
The integrated database of human OS, named HOsDb, aims to provide a high-quality collection of human OS-related genes, methylation sites, CNVs, miRNAs, TFs, and SNPs based on literature mining, microarray analysis, and database retrieval. The data collection and processing steps are illustrated in Fig. 1.

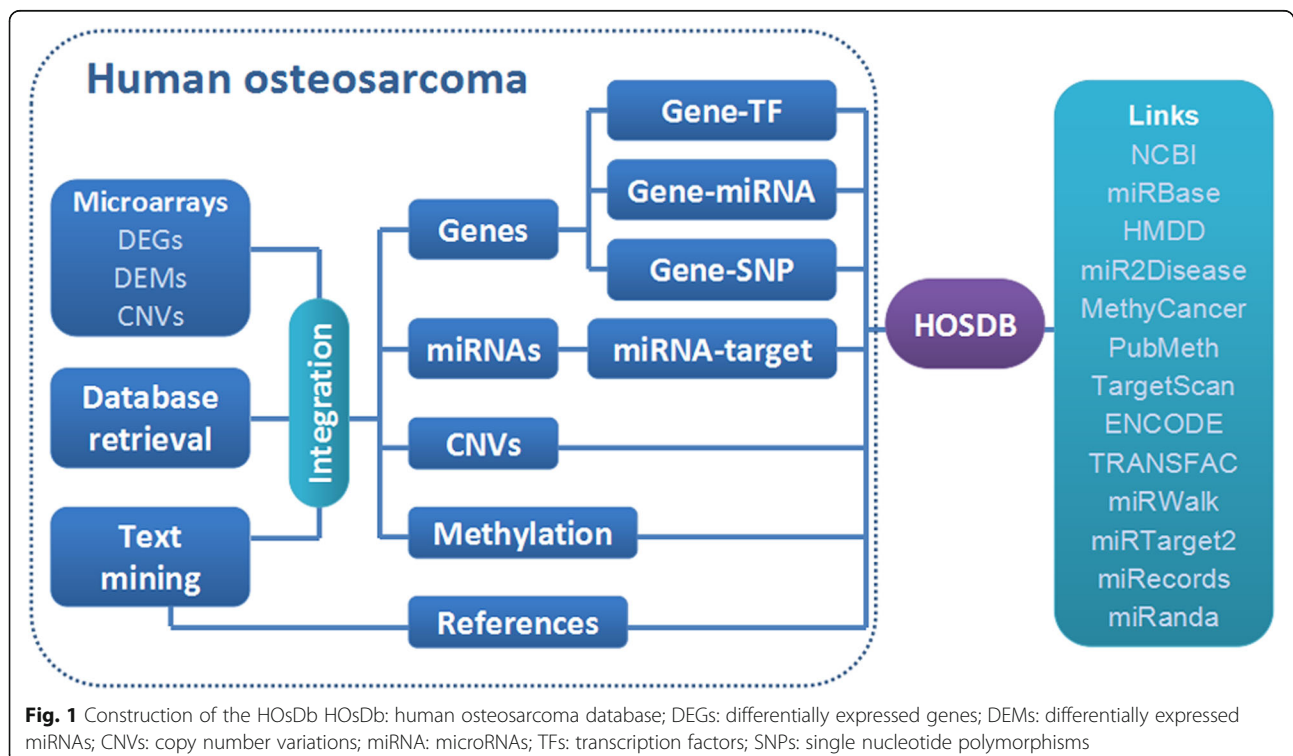**Fig. 1** Construction of the HOsDb HOsDb: human osteosarcoma database; DEGs: differentially expressed genes; DEMs: differentially expressed miRNAs; CNVs: copy number variations; miRNA: microRNAs; TFs: transcription factors; SNPs: single nucleotide polymorphisms

Table 1 Information of the included datasets

Category		GEO accession number	Platform	Sample type	Sample size	Experiment design
Gene	Tumor vs. normal	GSE11414	GPL6244	Cell lines	Tumor: 4; normal: 2	genome-wide comparison of gene expression and identified genes that are differentialy expressed in osteosarcoma (U2OS, MG63) cell lines relative to normal human osteoblasts (HOB)
		GSE12865	GPL6244	Tumor tissues and normal cell line	Tumor: 12; normal: 2	genome-wide comparison of gene expression and identified genes that are differentialy expressed in osteosarcoma tumour samples relative to normal human osteoblasts (HOB)
		GSE14359	GPL96	Tumor tissues and normal cell line	Tumor: 18; normal: 2	mRNA from 5 frozen conventional osteosarcoma and 4 osteosarcoma lung metastases tumor samples and mRNA from fresh primary osteoblast cells (HOB) were extracted and hybridized to HG U133A microarrays
		GSE16088	GPL96	Tumor and normal tissues, as well as tumor cell lines	Tumor: 17; normal: 6	Profiles of human osteosarcoma and three normal tissues, single channel design
		GSE19276	GPL6848	Tumor and normal tissues	Tumor: 44; normal: 5	Gene expression profiling of primary osteosarcoma biopsies and compared the results to gene expression profiling of non-malignant bone to identify differentially expressed genes unique to OS in the context of the bone micorenvironment
		GSE28424	GPL13376	Tumor cell lines and normal tissues	Tumor: 19; normal: 4	19 osteosarcoma cell lines, 4 normal bones used as controls. No replicates. The group of osteosarcomas are compared to the group of normal bones.
		GSE30807	GPL570	Tumor cell lines and normal bone mesenchymal stem cells	Tumor: 2; normal: 1	To analysis stem/progenitor cell-associated genes and molecules involved in regulation of self-renewal signaling pathways of cancer stem cells between UT2 cells and its parent cells: U2OS (MSC works as positive control here)
		GSE36001	GPL6102	Tumor cell lines and normal osteoblast and bone cells.	Tumor: 19; normal: 6	Comparison of gene expression patterns in 19 osteosarcoma cell lines and 6 normal samples (osteoblasts and bones)
		GSE42352	GPL10295	Tumor cell lines, pre-chemotherapy biopsies, osteoblasts, mesenchymal stem cells	Tumor: 103; normal: 15	Gene set analysis on previously published genome-wide gene expression data of osteosarcoma cell lines ($n = 19$) and osteosarcoma pre-chemotherapy biopsies ($n = 84$), and characterizing expression of the insulin-like growth factor receptor signaling pathways in human osteosarcoma as compared with osteoblasts and with the hypothesized progenitor cells of osteosarcoma - mesenchymal stem cells.
		GSE56001	GPL10558	Tumor cells and normal mesenchymal stem cells	Tumor: 3; normal: 9	Analysis of gene changes in different genes modulation in mesenchymal stem cells and compared to primary human osteosarcoma cells
		GSE9508	GPL6076	Tumor and normal biopsies	Tumor: 34; normal: 5	Two-colour experiment. 7 samples for non-metastatic patients, 6 of which are analyzed in duplicate (dye-swaps); 11 samples for metastatic patients, 10 of which are analyzed in duplicate (dye-swaps); 5 samples of non-malignant bone analyzed individualy, no dye-swaps (i.e. 5 biological replicates).
	Metastasis vs. non-metastasis	GSE14359	GPL96	Conventional osteosarcoma and lung metastases tumor samples	Metastasis: 8; non-metastasis: 10	mRNA from 5 frozen conventional osteosarcoma and 4 osteosarcoma lung metastases tumor samples and mRNA from fresh primary osteoblast cells (HOB) were extracted and hybridized to HG U133A microarrays
		GSE18947	GPL570	Low and high metastatic potential cell sublines	Low metastasis: 3; high metastasis: 3	The assay was performed among three pairs of cublines, the first two pairs of sublines comes from the different passage of sublines established with orthotopic transplantation

Sun *et al. BMC Cancer*   (2020) 20:390

Page 4 of 9

Table 1 Information of the included datasets *(Continued)*

Category	GEO accession number	Platform	Sample type	Sample size	Experiment design
					under the established cell line named Sosp-9607, the other pair was screened by the tail-vein injection method of commercial avaliable cell line-Saos-2.
	GSE21257	GPL10295	Metastatic and non-metastatic tumor biopsies	Metastasis: 34; non-metastasis: 19	Pre-chemotherapy biopsies of osteosarcoma patients who developed metastases within 5 yrs. ($n = 34$) were compared with pre-chemotherapy biopsies of osteosarcoma patients who did not develop metastases within 5 yrs. ($n = 19$)
	GSE9508	GPL6076	Metastatic and non-metastatic tumor biopsies	Metastasis: 21; non-metastasis: 13	Two-colour experiment. 7 samples for non-metastatic patients, 6 of which are analyzed in duplicate (dye-swaps); 11 samples for metastatic patients, 10 of which are analyzed in duplicate (dye-swaps); 5 samples of non-malignant bone analyzed individualy, no dye-swaps (i.e. 5 biological replicates).
Drug-treated vs. untreated	GSE16089	GPL570	Methotrexate-sensitive and –resistant Saos-2 cells	Methotrexate-sensitive samples: 3; methotrexate-resistant samples: 3	Two cell lines are compared, which are Saos-2 osteosarcoma cells sensitive to methotrexate and Saos-2 cells resistant to 10e-6 M methotrexate. Six samples are provided which correspond to triplicates of each cell line.
	GSE24401	GPL1456	Atorvastatin-treated and -untreated Saos-2 cells	Atorvastatin-treated samples: 3; atorvastatin-untreated samples: 3	Dye balance-experiment comparing atorvastatin treated Saos-2 cells versus untreated cells at 6, 15 and 24 h using 2 biological replicates
miRNA	GSE28423	GPL8227	Tumor cell lines and normal bones	Tumor cell lines: 19; normal bones: 4	19 osteosarcoma cell lines, 4 normal bones used as controls. No replicates. The group of osteosarcomas are compared to the group of normal bones.
CNV	GSE12830	GPL4091, GPL9128	Tumor tissues	20	Integrative whole-genome analysis of DNA copy number, promoter methylation and gene expression using 10 osteosarcomas with 2 biological replicates
	GSE7077	GPL2879	Four osteosarcoma-derived cell lines: U-2 OS, HOS, MG-63 and SAOS-2	4	To utilize oligonucleotide array CGH to identify microaberrations in osteosarcomas, likely to contain genes involved in osteosarcoma tumor oncogenesis. A better understanding of the underlying molecular genetic events leading to tumor initiation and progression could result in the identification of prognostic markers and therapeutic targets.
	GSE9654	GPL2879		10	To utilize oligonucleotide array CGH and FISH analysis to derive possible genomic signatures of chromosomal instability in osteosarcoma tumors

"-" in the column of "PubMed ID" means that there is no published study so far. *GEO* Gene Expression Omnibus, *CGH* Comparative genomic hybridization, *FISH* Fluorescence in situ hybridization

OS-related genes

Initially, mRNA expression microarrays related to OS were downloaded from the GEO database [19]. Detailed information regarding the datasets used, such as the GEO accession number and sample type and size, is shown in Table 1. The corresponding experimental conditions were tumor vs. normal, metastasis vs. non-metastasis, or drug-treated vs. untreated. Raw Affymetrix data in CEL file format were read using Affy [20] and normalized using the robust microarray analysis (RMA) method [21]. The downloaded normalized expression matrix was used for analysis of data generated using Illumina and Agilent platforms.

Differentially expressed genes (DEGs), defined as OS-related DEGs, were identified using the Linear Models for Microarray and RNA-Seq Data (limma) package [22] with a cut-off value of |log fold change (FC)| > 1 and false discovery rate (FDR) [23] < 0.05. A total of 6964 OS tumor-related, 685 OS metastasis-related, and 1589 OS drug-related DEGs were identified (Table 2). Literature mining of the PubMed collection was used to generate a list of known OS tumor-related and OS metastasis-related genes. A total of 505 genes related to OS tumor and 87 genes related to OS metastasis were found in the published literature. A list of OS-related genes was then collated by

Sun *et al. BMC Cancer*   (2020) 20:390

Page 5 of 9

Table 2 Results of data collection and analysis

	Tumor vs. normal	Metastasis vs. non-metastasis	Drug-treated vs. untreated
OS-related gene			
DEG (mRNA expression microarray)	6964	685	1589
Known gene (text mining)	505	87	–
Total OS-related gene	7191	763	1589
Gene-TF pair (database)	190,362	21,131	41,135
Gene-miRNA pair (database)	3,749,490	358,361	767,674
Gene-SNP (database)	28,386	2532	3943
OS-related miRNA			
DEM (miRNA expression microarray)	209	–	–
miRNA (database)	31	–	–
Total OS-related miRNA	240		
OS-related CNV			
CNV (CGH microarray)	1695	–	–
OS-related methylation			
Gene methylation (database)	18	–	–

OS Osteosarcoma, *DEG* Differentially expressed gene, *DEM* Differentially expressed miRNA, *miRNA* microRNA, *CNV* Copy number variation, *CGH* Comparative genomic hybridization

integrating OS-related DEGs identified by microarray analysis and OS-related genes identified by literature mining. Using this approach, 7191 OS tumor-related genes (Supplementary Table 1), 763 OS metastasis-related genes, and 1589 OS drug-related genes were identified (Table 2).

A list of TFs targeting OS-related genes was obtained from the TRANSFAC [24] and ENCODE databases [25]. We found 299 OS tumor-, 207 OS metastasis-, and 194 OS drug-related TFs, which corresponded to 190,362, 21,131, and 41,135 gene-TF pairs, respectively (Table 2). The miRNAs targeting OS-related genes were extracted from existing databases, including miRanda (Good mirSVR score part; release: August 2010) [26], miRecords (version 4) [27], miRTarget2 (version 4) [28], miRWalk (validated targets only) [29], and TargetScan (release 6.2) [30]. A total of 3,749,490, 358,361, and 767,674 gene-miRNA pairs related to OS tumor, metastasis, and drug treatments, respectively, were identified (Table 2). SNPs in OS-related genes were extracted from the National Center for Biotechnology Information (NCBI) dbSNP database (updated on 2014.05.29) [31]. We found 28,386, 2532, and 3943 SNPs in genes related to OS tumor, metastasis, and drug treatment, respectively (Table 2).

OS-related miRNAs
Normalized miRNA expression microarray data related to OS were also downloaded from the GEO database (Table 1). Differentially expressed miRNAs (DEMs) were identified using the limma package with a cutoff value of $|logFC| > 1$ and FDR < 0.05. Known OS-related miRNAs were extracted from the miR2Disease database (updated on 2011.04.14) [10] and HMDD database (updated on

2012.09.09) [11]. In total, 209 OS-related DEMs were identified based on miRNA expression microarray, and 31 known OS-related miRNAs were identified in the miR2-Disease and HMDD databases, generating a final count of 240 OS-related miRNAs for inclusion (Table 2).

OS-related CNVs
Normalized, comparative genomic hybridization (CGH) microarray data were downloaded from the GEO database (Table 1) and analyzed using DNAcopy [32] and cghMCR packages [33]. The criteria were set at (Segment Gain or Loss (> 0.2 and incidence > 30%. A total of 1695 genes with CNVs in OS were identified (Table 2).

OS-related methylation sites
MethyCancer [12] and PubMeth [34] databases were searched using the keyword "osteosarcoma." Eighteen genes with methylation sites related to OS were identified for further analysis (Table 2).

Data storage
The data obtained using the methods described were collated and used to construct the integrated human OS database (HOsDb), which is available for use at www.hosdatabase.com. HOsDb is a one-stop comprehensive platform for OS researchers.

Database description
The HOsDb is a search engine that can be used to search detailed information on each OS-related term stored in the database. Terms include 'Home,' 'Introduction,' 'Tumor vs. normal,' 'Metastasis vs. non,'

'Drug-treated vs. untreated,' 'miRNA,' 'copy number variation,' 'methylation,' 'Related database,' and 'Download.' The 'Tumor vs. normal,' 'Metastasis vs. non,' and 'Drug-treated vs untreated' terms on the home page focus on OS-related genes, as well as TFs, miRNAs and SNPs targeting OS-related genes. Users can query a gene symbol in the search bar located at the top of the linked pages. After inputting the gene symbol, all information related to that gene will be displayed in a new page, including gene/TF/miRNA/ SNP symbol, synonyms, full name, logFC, *p*-value, GEO microarray ID, gene/miRNA regulation direction in OS, miRNA targets, and links to publications in PubMed. To see more details about their gene of interest, users can click on the gene symbol link, and the NCBI page and results related to the gene of

interest will appear (Fig. 2). The 'miRNA' term links users to a list of OS-related miRNAs, and users can search a particular miRNA by inputting its symbol in the search bar. Notably, users can define their own thresholds (logFC and p-value) for gene or miRNA expression. However, the default settings are logFC > 1 and p-value < 0.05 (Fig. 3a). The 'copy number variation' term generates a list of genes with CNVs in OS. Users can query whether a certain gene undergoes changes in copy number in OS or not by inputting the corresponding gene ID or symbol (Fig. 3b). The 'methylation' term lists all genes with methylation sites related to OS. Users can input a gene symbol to check whether its sequence has methylation sites in OS or not (Fig. 3c). The 'Related database' terms include several internal resources or databases, which are cross-linked in HOsDb, including NCBI,

Fig. 2 Schematic diagram of the workflow for collating OS-related genes OS: osteosarcoma; HOsDb: human osteosarcoma database; TF: transcription factor; miRNA: microRNA; SNP: single nucleotide polymorphism; ID: identifier. 'Tumor vs. normal,' 'Metastasis vs. non,' and 'Drug-treated vs. untreated' sections on the homepage are all focused on OS-related genes

Fig. 3 Schematic diagram of the workflow for collating OS-related miRNAs, CNVs, and methylation sitesa) miRNAs b) CNVs c) methylation sites. OS: osteosarcoma; HOsDb: human osteosarcoma database; miRNAs: microRNAs; CNV: copy number variation; ID: identifier

Sun *et al. BMC Cancer*   (2020) 20:390

Page 8 of 9

miRBase, HMDD, miR2Disease, MethyCancer, PubMeth, TargetScan, ENCODE, TRANSFAC, miRWalk, miRTarget2, miRecords, and miRanda. The 'Download' term allows users to obtain detailed information regarding DEGs, DEMs, TFs, SNPs, and CNVs that was used for HOsDb construction.

Utility and discussion

Compared with a previously established OS database [14], the HOsDb provides more information. For example, our analyses of mRNA and miRNA expression microarrays, and CGH microarray provide a comprehensive list of candidate genes, miRNAs, and CNVs, which will assist users to navigate through the complexity of OS. Moreover, the HOsDb contains detailed gene regulation information, such as potential TF- and miRNA-gene pairs associated with OS, which is convenient for the identification of novel gene relationships involved in OS. Furthermore, information regarding SNPs in OS-related genes is provided in the HOsDb, which will help direct further studies of OS-related SNPs. The OS-related CNVs listed in the HOsDb were generated through analysis of three CGH microarray datasets. Thus, they are more reliable than those generated from a single dataset. Additionally, the HOsDb incorporates a user-friendly interface, which makes all the features easily accessible.

Although data in the HOsDb were collected using a number of different platforms and approaches, all data were normalized prior to analysis, thus adding to the reliability of our results. However, microarray data regarding OS are likely to be constantly updated in the GEO database and next-generation sequencing studies can also provide OS-related data, which will provide new insights into OS biology. This updated information will need to be added to HOsDb, once it is available. Although the HOsDb has advantages over the only other known OS-related database in its current form, we plan to update the database periodically to consistently maintain the quality of OS-related data available, and thus, keep up to date with changes and improvements in the field.

Conclusions

The HOsDb provides a one-stop, comprehensive platform for human OS research that is quick and easily accessible. We believe that the HOsDb will be particularly attractive to communities and researchers interested in OS, and that the HOsDb will considerably facilitate research regarding the pathogenesis of OS.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12885-020-06719-2.

Additional file 1.

Abbreviations
CGH: Comparative genomic hybridization; CNV: Copy number variations; DEG: Differentially expressed gene; DEM: Differentially expressed miRNA; FDR: False discovery rate; GEO: Gene Expression Omnibus; HOsDb: human OS database; miRNA: microRNA; OS: Osteosarcoma; SNP: Single nucleotide polymorphism; TF: Transcription factor

Author contributions
Conception and design: Yifu Sun and Lishan Wang; collection and assembly of data: Changkuan Li and Rui Gu; data analysis and interpretation: Weidong Zang, Wei Song and Peng Xia; article writing: all authors; final approval of article: all authors.

Availability of data and materials
The datasets generated and analyzed during the current study are available in the HOsDb (www.hosdatabase.com).

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Orthopedics, China-Japan Union Hospital of Jilin University, Changchun 130033, P.R. China. [2]Eryun (Shanghai) Information Technology Co., Ltd, Shanghai 200241, P.R. China. [3]Department of Orthopedics, The Second Hospital of Jilin University, No.218 Ziqiang Street, Changchun 130022, China.

References
1. Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. Nat Rev Cancer. 2014;14(11):722.
2. Ritter J, Bielack S. Osteosarcoma. Ann Oncol. 2010;21(suppl_7):vii320–5.
3. Stiller C, Bielack S, Jundt G, Steliarova-Foucher E. Bone tumours in European children and adolescents, 1978–1997. Report from the automated childhood Cancer information system project. Eur J Cancer. 2006;42(13):2124–35.
4. Lindsey BA, Markel JE, Kleinerman ES. Osteosarcoma overview. Rheumatol Ther. 2017;4(1):25–43.
5. Fletcher CD, Unni KK, Mertens F. Pathology and genetics of tumours of soft tissue and bone, vol. 4: Iarc; 2002.
6. Porter D, Holden S, Steel C, Cohen B, Wallace M, Reid R. A significant proportion of patients with osteosarcoma may belong to Li-Fraumeni cancer families. J Bone Joint Surg Br Vol. 1992;74(6):883–6.
7. Hayden JB, Hoang BH. Osteosarcoma: basic science and clinical implications. Orthop Clin N Am. 2006;37(1):1–7.
8. Amberger J, Bocchini CA, Scott AF, Hamosh A. McKusick's online Mendelian inheritance in man (OMIM®). Nucleic Acids Res. 2009;37(suppl 1):D793–6.
9. Maier H, Döhr S, Grote K, O'keeffe S, Werner T, de Angelis MH, Schneider R. LitMiner and WikiGene: identifying problem-related key players of gene regulation using publication abstracts. Nucleic Acids Res. 2005;33(suppl_2):W779–82.
10. Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, Li M, Wang G, Liu Y. miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Res. 2009;37(suppl 1):D98–D104.

Sun *et al. BMC Cancer*   (2020) 20:390

Page 9 of 9

11. Li Y, Qiu C, Tu J, Geng B, Yang J, Jiang T, Cui Q. HMDD v2. 0: a database for experimentally supported human microRNA and disease associations. Nucleic Acids Res. 2013;42(D1):D1070.

12. He X, Chang S, Zhang J, Zhao Q, Xiang H, Kusonmano K, Yang L, Sun ZS, Yang H, Wang J. MethyCancer: the database of human DNA methylation and cancer. Nucleic Acids Res. 2008;36(suppl 1):D836–41.

13. Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Prüß M, Reuter I, Schacherer F. TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res. 2000;28(1):316–9.

14. Poos K, Smida J, Nathrath M, Maugg D, Baumhoer D, Neumann A, Korsching E. Structuring osteosarcoma knowledge: an osteosarcoma-gene association database based on literature mining and manual annotation. Database (Oxford). 2014;2014:1–9.

15. Sadikovic B, Yoshimoto M, Chilton-MacNeill S, Thorner P, Squire JA, Zielenska M. Identification of interactive networks of gene expression associated with osteosarcoma oncogenesis by integrated molecular profiling. Hum Mol Genet. 2009;18(11):1962–75.

16. Walter I, Wolfesberger B, Miller I, Mair G, Burger S, Galle B, Steinborn R. Human osteosarcoma cells respond to sorafenib chemotherapy by downregulation of the tumor progression factors S100A4, CXCR4 and the oncogene FOS. Oncol Rep. 2014;31(3):1147–56.

17. Buddingh EP, Kuijjer ML, Duim RA, Burger H, Agelopoulos K, Myklebost O, Serra M, Mertens F, Hogendoorn PC, Lankester AC, et al. Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. Clin Cancer Res. 2011;17(8):2110–9.

18. Namlos HM, Meza-Zepeda LA, Baroy T, Ostensen IH, Kresse SH, Kuijjer ML, Serra M, Burger H, Cleton-Jansen AM, Myklebost O. Modulation of the osteosarcoma expression phenotype by microRNAs. PLoS One. 2012;7(10):e48086.

19. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207–10.

20. Gautier L, Cope L, Bolstad BM. Irizarry RA: affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20(3):307–15.

21. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. 2003;4(2):249–64.

22. Smyth GK. Limma: linear models for microarray data. In: Bioinformatics and computational biology solutions using R and Bioconductor. NewYork: Springer; 2005. p. 397–420.

23. Storey JD. False discovery rate. In: International encyclopedia of statistical scienc. NewYork: Springer; 2011. p. 504–8.

24. Matys V, Fricke E, Geffers R, Gößling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV. TRANSFAC®: transcriptional regulation, from patterns to profiles. Nucleic Acids Res. 2003;31(1):374–8.

25. Consortium EP. The ENCODE (ENCyclopedia of DNA elements) project. Science. 2004;306(5696):636–40.

26. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human microRNA targets. PLoS Biol. 2004;2(11):e363.

27. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA–target interactions. Nucleic Acids Res. 2009;37(suppl 1):D105–10.

28. Wang X, El Naqa IM. Prediction of both conserved and nonconserved microRNA targets in animals. Bioinformatics. 2008;24(3):325–32.

29. Dweep H, Sticht C, Pandey P, Gretz N. miRWalk–database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform. 2011;44(5):839–47.

30. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005;120(1):15–20.

31. Sherry S, Ward M-H, Kholodov M, Baker J, Phan L, Smigielski E, Sirotkin K. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29(1):308.

32. Seshan VE, Olshen AB. DNAcopy: a package for analyzing DNA copy data. R package version; 2013.

33. Zhang J, Feng B, Zhang MJ. biocViews microarray C: package 'cghMCR'. R package version; 2013.

34. Ongenaert M, Van Neste L, De Meyer T, Menschaert G, Bekaert S, Van Criekinge W. PubMeth: a cancer methylation database combining text-mining and expert annotation. Nucleic Acids Res. 2007;36(suppl_1):D842–6.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.