RESEARCH ARTICLE

A five metastasis-related long noncoding RNA risk signature for osteosarcoma survival prediction

SiYuan Yu[1,3], FengLing Shao[2,3], HuiJun Liu[4] and QingQing Liu[2,3*]

Abstract

Background: Osteosarcoma is a highly malignant and common bone tumour with an aggressive disease course and a poor prognosis. Previous studies have demonstrated the relationship between long noncoding RNAs (lncRNAs) and tumorigenesis, metastasis, and progression.

Methods: We utilized a large cohort from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database osteosarcoma project to identify potential lncRNAs related to the overall survival of patients with osteosarcoma by using univariate and multivariate Cox proportional hazards regression analyses. Kaplan–Meier curves were generated to evaluate the overall survival difference between patients in the high-risk group and the low-risk group. A time-dependent receiver operating characteristic curve (ROC) was employed, and the area under the curve (AUC) of ROC was measured to assess the sensitivity and specificity of the multi-lncRNA signature.

Results: Five lncRNAs (RP11-128N14.5, RP11-231|13.2, RP5-894D12.4, LAMA5-AS1, RP11-346L1.2) were identified, and a five-lncRNA signature was constructed. The AUC for predicting 5-year survival was 0.745, which suggested good performance of the five-lncRNA signature. In addition, functional enrichment analysis of the five-lncRNA-correlated protein-coding genes (PCGs) was performed to show the biological function of the five lncRNAs. Additionally, PPI network suggested RTP1 is a potential biomarker that regulates the prognosis of osteosarcoma.

Conclusions: We developed a five-lncRNA signature as a potential prognostic indicator for osteosarcoma.

Keywords: Long noncoding RNA, Bioinformatics analysis, Prognostic biomarker, Osteosarcoma, TARGET database

Background

Osteosarcoma is a highly malignant and common bone tumour with an aggressive and poor prognosis and is prevalent in adolescents and young adults [1, 2]. Approximately 80% of osteosarcomas occur in the extremities, most commonly in the metaphysis of the long bone around the knee joint, and the other 20% occur in the axial bone and pelvis [3]. In addition, metastasis, especially lung metastasis, is an important factor in determining the prognosis and survival of osteosarcoma. It has been reported that high-dose methotrexate, vincristine, and folinic acid could improve the survival rate of patients. However, the survival rate of patients with metastatic osteosarcoma remains unchanged [4]. Therefore, we aimed to search for prognostic targets of osteosarcoma.

The function of long noncoding RNAs (lncRNAs) in many areas has gradually gained an appreciation and has been extensively studied. lncRNAs are defined as transcripts of more than 200 nucleotides that are not translated into proteins [5] and can participate in transcription,

*Correspondence: 18366116578@163.com
[2] Department of Pediatric Surgical Oncology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, China
Full list of author information is available at the end of the article

Yu *et al. BMC Med Genomics*   (2021) 14:124

Page 2 of 10

post-transcriptional regulation, and protein regulation. Findings from numerous epidemiological studies have linked lncRNAs to tumours, which suggests that lncRNAs play an important role as biomarkers [6, 7] and therapeutic targets [8, 9]. In addition, the role of lncRNAs in osteosarcoma is currently of significant interest to many researchers. Previous studies have demonstrated the relationship between lncRNAs and tumorigenesis [9, 10], metastasis [11], progression [12], and chemotherapeutic drug resistance [13] in osteosarcoma. Unfortunately, there is limited intervention research on the impact of potential prognostic lncRNA signatures for osteosarcoma.

The Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database provides RNA-sequencing (mRNA-seq) data on several types of cancers, including acute lymphoblastic leukaemia, myeloid leukaemia, kidney tumours, neuroblastoma, and osteosarcoma. In this study, a large cohort from the TARGET osteosarcoma project was used to identify potential lncRNAs related to the overall survival of patients with osteosarcoma. In addition, we identified a five-lncRNA signature that could predict the prognosis of patients with osteosarcoma. Our study suggests the potential functions of lncRNAs in osteosarcoma.

Methods
Data collection and preprocessing
Data of mRNA-seq and relevant clinical sample data of osteosarcoma patients were collected through the TARGET data matrix (https://ocg.cancer.gov/programs/target). We downloaded data for 306 clinical samples in total, but only 102 clinical samples had mRNA-seq data, and after deleting duplicate data, there were only 97 samples. Altogether, 24 metastatic samples and 73 non-metastatic samples were used to determine differentially expressed lncRNAs by using the edger package in R software (4.0.2). Differently expressed lncRNAs were selected based on threshold log foldchange (FC) > 1, adjusted $p < 0.05$.

Identification of a prognostic lncRNA signature
We excluded patients without vital status and lncRNA sequence data. A total of 95 patients (providing 23 metastatic samples and 72 non-metastatic samples) were concluded in the research. Then, univariate Cox proportional hazards regression analysis was used to select significant lncRNAs related to the survival status (Additional file 1: Table S1). Finally, we identified 5 lncRNAs (RP11-128N14.5, RP11-231|13.2, RP5-894D12.4, LAMA5-AS1, RP11-346L1.2) according to $P < 0.05$ and HR > 1 and using R software (4.0.2).

Independence of the multi-lncRNA signature for predicting the prognosis of patients with osteosarcoma
To construct the prognostic multi-lncRNA signature, we conducted multivariate Cox proportional hazards regression analysis. A risk score was calculated by integrating the lncRNA expression profiles and relative multivariate Cox regression coefficients (Additional file 2: Table S2). According to the risk score, patients with osteosarcoma consisted of a group with low-risk and a group with high-risk. Then, Kaplan−Meier curves were generated to estimate the whole surviving discrepancy among invalids in the team with high-risk and team with low-risk. In addition, a receiver that was time-oriented and operated characteristic curve (ROC) was adopted to identify specificity and sensitivity of multi-lncRNA signature with the help of measuring region under the curve of ROC [14, 15].

Functional enrichment analysis
lncRNA-correlated protein-coding genes (PCGs) was the base of the analysis of functional enrichment. A Pearson analysis of connection for every prognostic lncRNAs of PCGs according to sequence data of TARGET data matrix (Additional file 3: Table S3) was launched. PCGs about Pearson coefficients > 0.5 were reckoned to be lncRNA-related PCGs. Next, the pathway analysis of gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) was operated through the database aimed for annotation, visualization, and integrated discovery (http://metascape.org/) Internet tool [16]. Biological process (BP) terms and KEGG pathways of $P < 0.05$ were thought to be important.

Protein–protein interaction network construction
Protein–Protein interaction (PPI) analysis of lncRNA-correlated PCGs was constructed on the STRING protein database 11.0 (http://string-db.org/). We set the interaction score > 0.4 as the cut-off criterion[17, 18].

Table 1 Clinicopathologic characteristics of patients with osteosarcoma

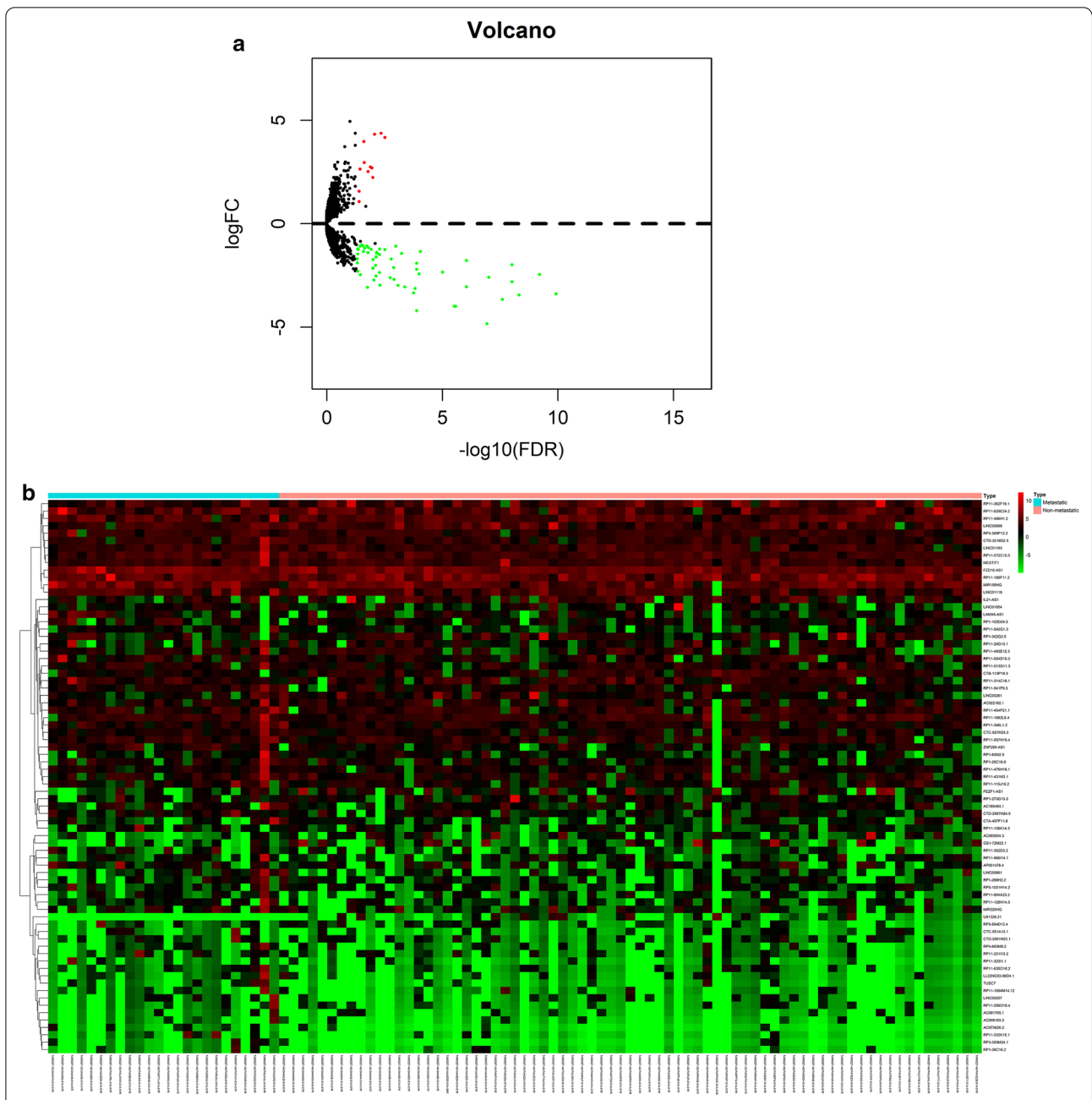| Characteristics	Sample size	Ratio (%)
Sex		
Female	39	41.1
Male	56	58.9
Age, y		
≤18	73	76.8
>18	22	23.2
Disease state at diagnosis		
Metastatic	23	24.2
Non-metastatic	72	75.8
Vital status		
Alive	57	60
Dead	38	40

Fig. 1 **a** Volcano plot of differentially expressed long noncoding RNAs (lncRNAs) between metastatic osteosarcoma samples and non-metastatic osteosarcoma samples; a total of 75 lncRNAs differentially expressed lncRNAs were identified, including 63 downregulated lncRNAs and 12 upregulated lncRNAs (cut-off criteria: log foldchange(FC) > 1, adjusted $p < 0.05$). Red represents upregulated lncRNAs, and green represents down-regulated lncRNAs. **b** Hierarchical clustering analysis of differentially expressed lncRNAs. Red represents upregulated lncRNAs, and green represents down-regulated lncRNAs. LncRNA, long noncoding RNA

Results

Resource identification initiative statement

We downloaded 306 clinical samples, but only 102 clinical samples had mRNA-seq data, and after deleting duplicate data, only 97 samples remained. After excluding patients without vital status and lncRNA sequence data, 95 patients (providing 23 metastatic samples and 72 non-metastatic samples) were included in this study. Specific information is shown in Table 1.

Yu *et al. BMC Med Genomics*   (2021) 14:124

Page 4 of 10

Table 2 Multivariate Cox proportional hazards model of the five-lncRNA signature

Gene symbol	Coef	Exp (coef)	SE (coef)	HR	Pr(>|z|)
RP11-128N14.5	0.247734	1.281119	0.130509	1.340221	0.057669
RP11-231|13.2	0.287661	1.333306	0.182852	1.499192	0.115675
RP5-894D12.4	0.294636	1.342637	0.129259	1.470435	0.022642
LAMA5-AS1	0.310018	1.36345	0.11846	1.235036	0.008869
RP11-346L1.2	0.232996	1.262377	0.133438	1.379240	0.080793

Identification of differentially expressed lncRNAs

Since differential analysis of expression, about 75 different meaning lncRNAs in total were determined, for example, 63 downregulated lncRNAs and 12 upregulated lncRNAs (cut-off criteria: $logFC > 1$, adjusted $p < 0.05$). Volcano plot analysis was used to assess variation in lncRNA expression between metastatic samples and non-metastatic samples (Fig. 1a). In addition, hierarchical clustering was used to distinguish these two populations based on lncRNA expression data (Fig. 1b). Then, we identified 5 lncRNAs (RP11-128N14.5, RP11-231|13.2, RP5-894D12.4, LAMA5-AS1, RP11-346L1.2) according to $P < 0.05$ and $HR > 1$ by using univariate Cox proportional hazards regression analysis.

Results of the multivariate Cox proportional hazards model

The multivariate Cox proportional hazards model was established with 5 lncRNAs (RP11-128N14.5, RP11-231|13.2, RP5-894D12.4, LAMA5-AS1, RP11-346L1.2). Multivariate Cox regression analysis was operated on the basis of lncRNA expression and patient survival (Table 2). The expression of RP11-128N14.5, RP11-231|13.2, RP5-894D12.4, LAMA5-AS1, and RP11-346L1.2 was actively connected with survival risk. Osteosarcoma patients were divided into high- and low-risk patient groups according to the score of risk (Fig. 2a). A heatmap of the meaning of the 5 lncRNAs is displayed in Fig. 2b. With the risk score increasing, the expression of RP11-128N14.5, RP11-231|13.2, RP5-894D12.4, LAMA5-AS1, and RP11-346L1.2 increased. The connection between the survival of osteosarcoma patients and the risk score (Fig. 2c) suggested that as the score of risk rose, the risk of death in osteosarcoma patients increased, indicating that high-risk scores osteosarcoma patients have shorter surviving times than low-risk scores patients.

Survival prediction in osteosarcoma patients using the five lncRNAs

The multivariate Cox proportional hazards model was built on the basis of the expression of 5 lncRNAs and the survived osteosarcoma patients. The osteosarcoma invalids consisted of high- or low-risk patients according to the score of risk. Next, Kaplan–Meier analysis was utilized to detect the connection between the risk score and patient five-year survival (Fig. 3a). These outcome suggested that the survival rate of high-risk patients was obviously lower than that in with low-risk patients ($p = 9.221e−03$). In addition, we constructed a ROC curve to assess the success of the binomial classification model of high- and low-risk patients (Fig. 3b). The AUC value was 0.745, which suggests that patients can consist of high-risk group and low-risk group, with clear discrepancy in survival, according to risk score. LncRNAs RP11-128N14.5, RP11-231|13.2, RP5-894D12.4, LAMA5-AS1, and RP11-346L1.2 have the potential to estimate the survival of patients with osteosarcoma.
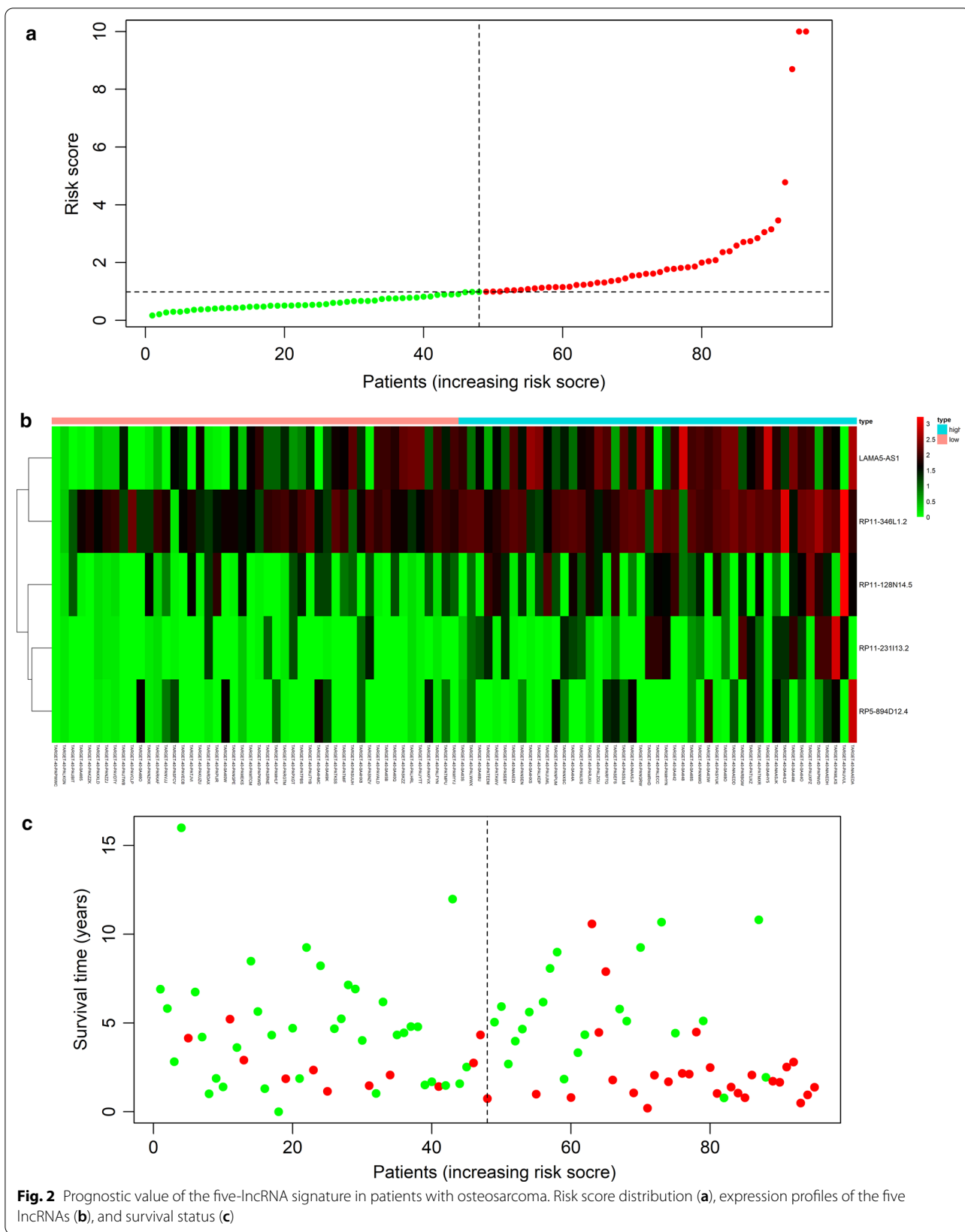
Functional enrichment analysis

950 related PCGs of a Pearson coefficient greater than 0.5 was identified. Final outcome of analysis of functional enrichment showed that the PCGs were mostly rich in BP terms, for example, multicellular organismal process, response to stimulus, reproductive processes, localization, cellular processes, and metabolic processes (Fig. 4a). The KEGG pathway analysis results showed that olfactory transduction, steroid hormone biosynthesis, neuroactive ligand-receptor interaction, and arachidonic acid metabolism were the main associated pathways (Fig. 4b).
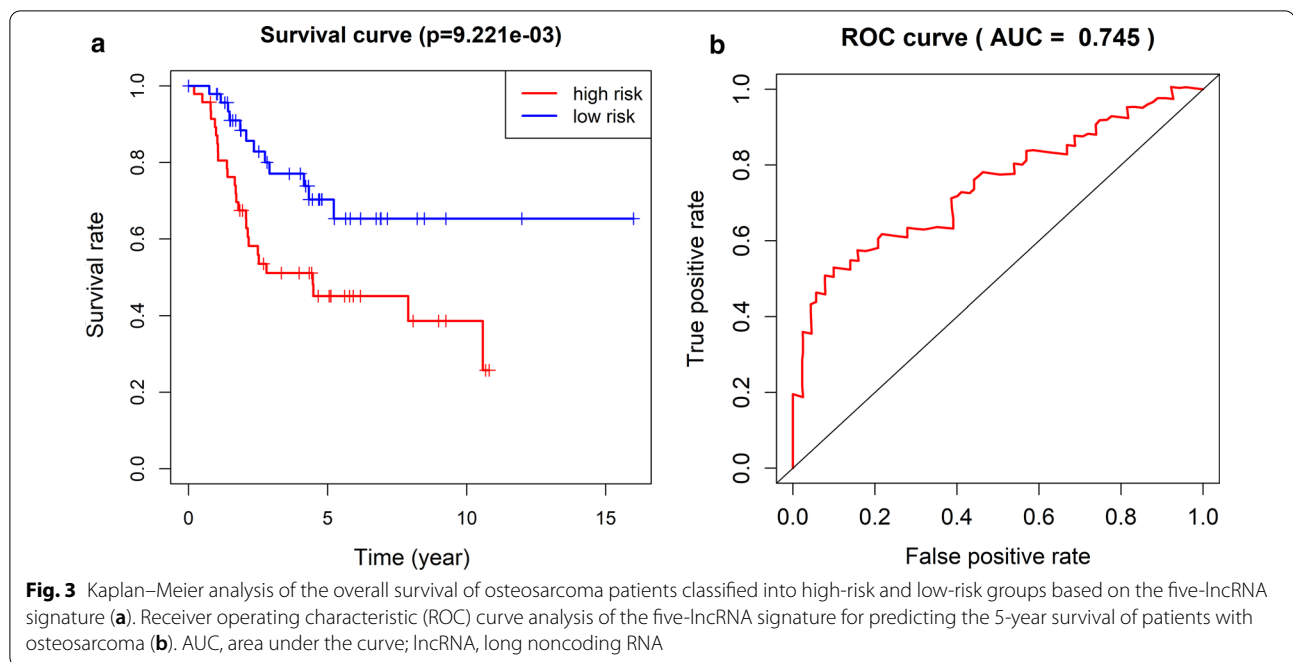
Protein–protein interaction network construction

The PPI network of lncRNA-correlated PCGs was also employed with the STRING protein database 11.0 (Fig. 5a). The number of nodes was 270, and the number of edges was 663. The average node degree was 4.91 and the average local clustering coefficient was 0.719. We also counted the number of times that each node appeared in the network (Fig. 5b). Finally identified that RTP1 may play a central role in the PPI network.

Discussion

Osteosarcoma is the most common primary bone tumour, and the five-year survival rate of patients with metastasis is much lower than that of patients without metastasis [19]. Yuan Li et al. [20] demonstrated that the expression of four genes, HIPK2, MAP3K5, CK5, and KRT5, was correlated with survival risk by using a multivariate Cox proportional hazards model. These results show that KRT5, HIPK2, MAP3K5, and CD5 serve as prognostic factors in osteosarcoma patients.

Fig. 2 Prognostic value of the five-lncRNA signature in patients with osteosarcoma. Risk score distribution (**a**), expression profiles of the five lncRNAs (**b**), and survival status (**c**)

Yu *et al. BMC Med Genomics*   (2021) 14:124

Page 6 of 10

Fig. 3 Kaplan–Meier analysis of the overall survival of osteosarcoma patients classified into high-risk and low-risk groups based on the five-lncRNA signature (**a**). Receiver operating characteristic (ROC) curve analysis of the five-lncRNA signature for predicting the 5-year survival of patients with osteosarcoma (**b**). AUC, area under the curve; lncRNA, long noncoding RNA

In addition, lncRNAs have emerged as potential prognostic biomarkers of tumours. Although many studies have elucidated the relationship between lncRNAs and osteosarcoma [21–23], there is still a lack of knowledge regarding lncRNAs related to survival prediction in patients with osteosarcoma.
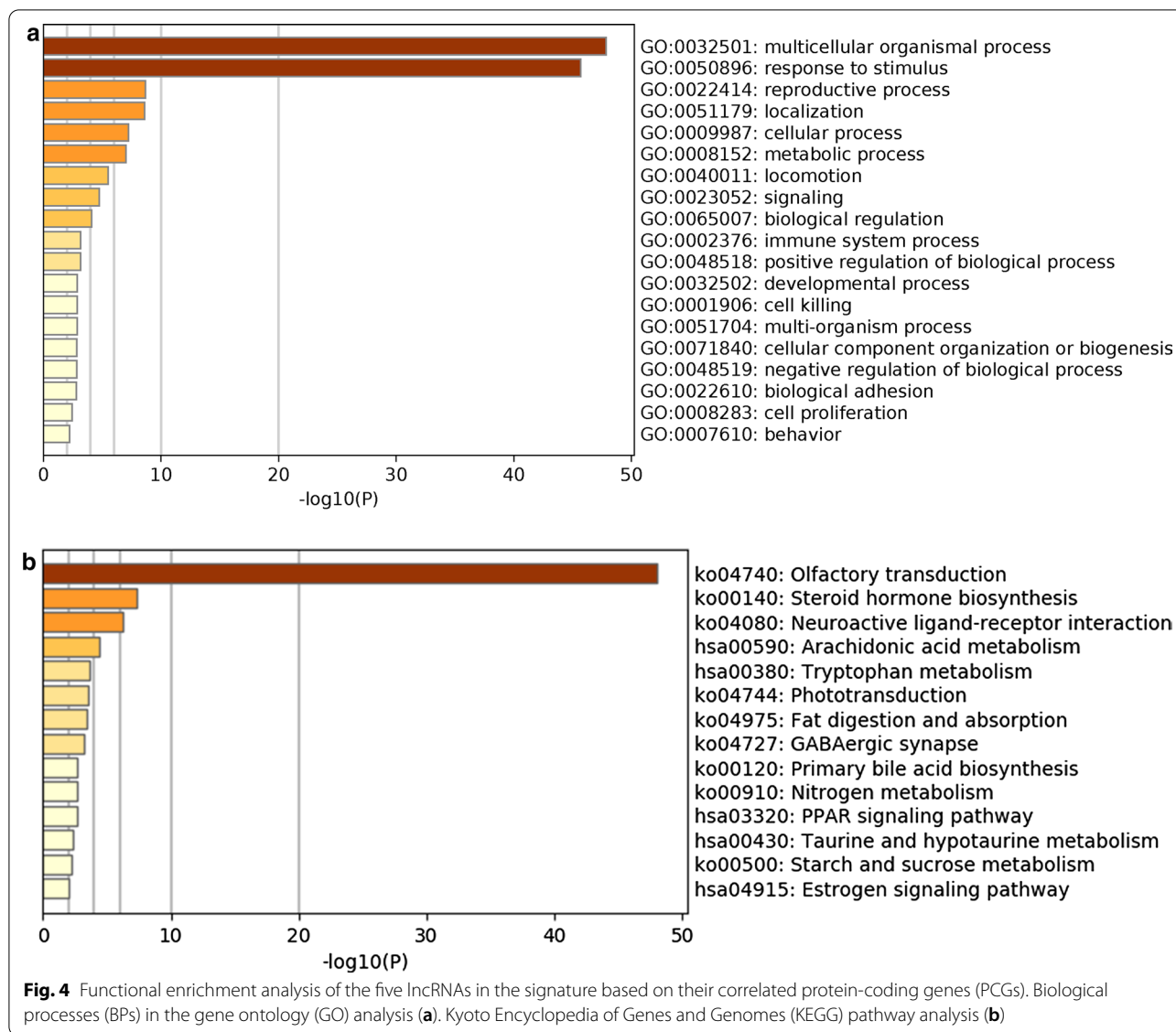
In this study, we utilized 95 osteosarcoma samples obtained from the TARGET database that were divided into metastatic and non-metastatic groups, and we analyzed their lncRNA expression profiles. Then, we constructed a lncRNA-based prognostic signature that accounts for the expression of five lncRNAs (RP11-128N14.5, RP11-231|13.2, RP5-894D12.4, LAMA5-AS1, and RP11-346L1.2), which were assumed to be significant according to their multivariate Cox regression coefficients. Unfortunately, Kaplan–Meier analysis revealed that the individual lncRNAs were not independent prognostic factors (results not shown in this article). However, according to the risk score that was calculated via the five-lncRNA prognostic signature, the samples could be divided into a high-risk group and a low-risk group. In addition, the Kaplan–Meier analysis suggested significant differences in overall survival between the two groups.

Among the five lncRNAs, RP11-128N14.5, RP11-231|13.2, and RP5-894D12.4 are sense intronic lncRNAs; RP11-346L1.2 is a lincRNA, and LAMA5-AS1 is an antisense lncRNA. RP11-128N14.5 has also been detected

in non-alcoholic fatty liver disease, and scientists have demonstrated that RP11-128N14.5 is upregulated at the intracellular, extracellular, and histological levels in non-alcoholic fatty liver disease [24]. Shen et al. [25] demonstrated that LAMA5-AS1 is related to the overall survival rate of patients with multiple myeloma. These research findings suggest that RP11-128N14.5 and LAMA5-AS1 are potential prognostic biomarkers for tumours. Our study also demonstrated that these lncRNAs show a close relationship with tumour prognosis.

Although it is well known that immunology regulations may affect cancer development [26], these five lncRNAs associated with osteosarcoma prognosis are rarely studied in immunodeficiency. In addition, the mRNA modifications are potentially new insights into biological basis [27, 28], especially N4-Acetylcytidine on RNA expression [29], these is no studies exploring the role of these five signatures in mRNA modification, especially N4-Acetylcytidine. Therefore, further researches correlated with immunodeficiency and mRNA modification are needed.

Next, to identify the functions of these five lncRNAs, we analyzed their related PCGs to explore the function of these lncRNAs according to the theory that co-expressed genes are more likely to be functionally related [30]. According to GO functional enrichment analysis, the BPs of the five-lncRNA signature mainly involved multicellular organismal processes, the response to stimulus, reproductive processes, localization, cellular processes,
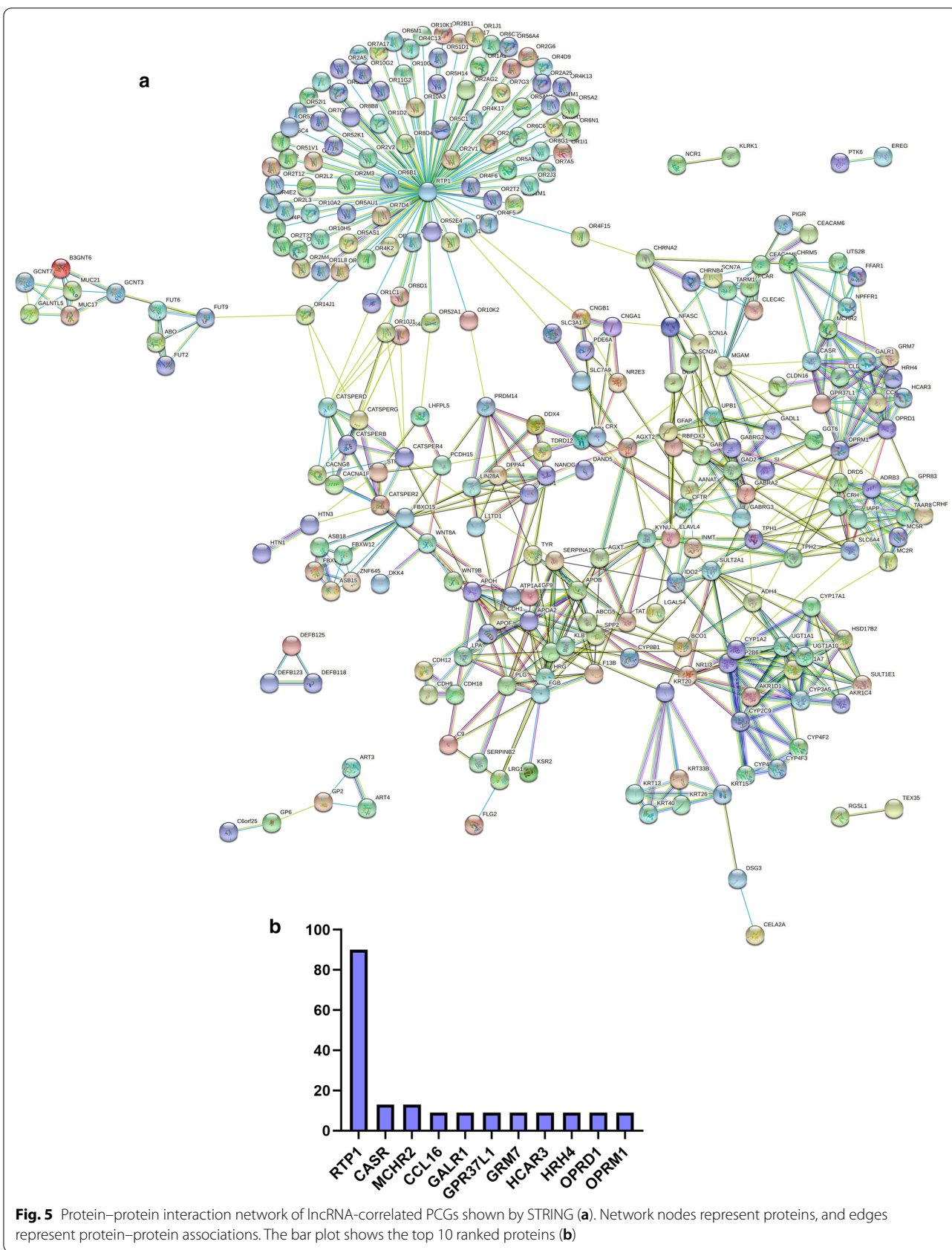
Yu *et al. BMC Med Genomics*   (2021) 14:124

Page 7 of 10

Fig. 4 Functional enrichment analysis of the five lncRNAs in the signature based on their correlated protein-coding genes (PCGs). Biological processes (BPs) in the gene ontology (GO) analysis (**a**). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis (**b**)

and metabolic processes. The KEGG pathway results suggested that olfactory transduction, steroid hormone biosynthesis, neuroactive ligand-receptor interaction, and arachidonic acid metabolism were related to the five-lncRNA signature. The relation of many PCGs with olfactory transduction in the KEGG pathway results was intriguing. Olfactory transduction has been reported in many different kinds of tumours, such as breast cancer [31], thyroid carcinoma [32], and hepatocellular carcinoma [33]. However, the relationship between olfactory transduction and osteosarcoma has not been reported, which provides a new point for osteosarcoma research.

Finally, PPI network was performed to explore the potential biological interaction or co-expression of lncRNA-correlated PCGs [34, 35]. We found RTP1 played an important role in the PPI network, suggesting RTP1 is a potential biomarker that regulates the prognosis of osteosarcoma. To our knowledge, there is rarely study reveal the mechanism of RTP1 in osteosarcoma, which indicated a new target for diagnosis or treatment of osteosarcoma.

Our study still has some limitations. Firstly, we only conducted high-throughput mRNA-seq profiling data analysis; therefore, for deeply understanding the mechanism, it is necessary for further functional verification of the importance of these five lncRNAs in the prognosis of osteosarcoma. Secondly, we only analyzed the data from the TARGET database with limited sample size, therefore other related omic data should be further collected to identify more significant biomarkers that could be interacting with the identified lncRNAs to jointly affect the

Yu *et al. BMC Med Genomics* (2021) 14:124

Page 8 of 10

Fig. 5 Protein–protein interaction network of lncRNA-correlated PCGs shown by STRING (**a**). Network nodes represent proteins, and edges represent protein–protein associations. The bar plot shows the top 10 ranked proteins (**b**)

Yu *et al. BMC Med Genomics*   (2021) 14:124

Page 9 of 10

survival status in a multilayer genetic regulation network [34–36]. Lastly, this study did not explore the possibility to construct a machine-learning tool based on these lncRNAs for prediction of survival due to technical reasons, and we will further try to use the biomarkers to predict the survival based on machine learning tools for better osteosarcoma survival prediction [37, 38].

Conclusions

In summary, we performed univariate and multivariate Cox regression analyses to identify five lncRNAs related to the survival of patients with osteosarcoma.

Abbreviations

LncRNAs: Long noncoding RNAs; TARGET: Therapeutically applicable research to generate effective treatments; ROC: Receiver operating characteristic curve; AUC: Area under the curve; PCGs: Protein-coding genes; mRNA-seq: RNA-sequencing; FC: Fold change; GO: Gene ontology; KEGG: Kyoto encyclopedia of genes and genomes; BP: Biological process; PPI: Protein–protein interaction.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12920-021-00972-5.

Additional file 1. Univariate Cox proportional hazards regression analysis of differentially expressed lncRNAs.

Additional file 2. Risk score of osteosarcoma patients based on five lncRNAs signature.

Additional file 3. Protein-coding genes (PCGs) correlated with five lncRNAs.

Authors' contributions
QL designed the study and wrote this article. SY downloaded data from the TARGET database and analyzed data. FS and HL interpreted the data. All authors read and approved the final manuscript.

Availability of data and materials
The datasets were downloaded from the TARGET data matrix (https://target-data.nci.nih.gov/Public/OS/mRNA-seq/).

Declarations

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 Anesthesiology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, China. [2]Department of Pediatric Surgical Oncology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, China. [3]Chongqing Key Laboratory of Pediatrics, Chongqing, China. [4]Reproductive Medical Center of Gansu Provincial Maternal and Child-Care Hospital, Lanzhou, China.

References

1. Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. Nat Rev Cancer. 2014;14(11):722–35.
2. Lu KH, Lu EW, Lin CW, Yang JS, Yang SF. New insights into molecular and cellular mechanisms of zoledronate in human osteosarcoma. Pharmacol Ther. 2020;214:107611.
3. Lamoureux F, Trichet V, Chipoy C, Blanchard F, Gouin F, Redini F. Recent advances in the management of osteosarcoma and forthcoming therapeutic strategies. Expert Rev Anticancer Ther. 2007;7(2):169–81.
4. Bernthal NM, Federman N, Eilber FR, Nelson SD, Eckardt JJ, Eilber FC, Tap WD. Long-term results (>25 years) of a randomized, prospective clinical trial evaluating chemotherapy in patients with high-grade, operable osteosarcoma. Cancer. 2012;118(23):5888–93.
5. Kopp F, Mendell JT. Functional classification and experimental dissection of long noncoding RNAs. Cell. 2018;172(3):393–407.
6. Zhang DY, Sun QC, Zou XJ, Song Y, Li WW, Guo ZQ, Liu SS, Liu L, Wu DH. Long noncoding RNA UPK1A-AS1 indicates poor prognosis of hepatocellular carcinoma and promotes cell proliferation through interaction with EZH2. J Exp Clin Cancer Res. 2020;39(1):229.
7. Shigeyasu K, Toden S, Ozawa T, Matsuyama T, Nagasaka T, Ishikawa T, Sahoo D, Ghosh P, Uetake H, Fujiwara T, et al. The PVT1 lncRNA is a novel epigenetic enhancer of MYC, and a promising risk-stratification biomarker in colorectal cancer. Mol Cancer. 2020;19(1):155.
8. Sun L, Zhu W, Zhao P, Wang Q, Fan B, Zhu Y, Lu Y, Chen Q, Zhang J, Zhang F. Long noncoding RNA UCA1 from hypoxia-conditioned hMSC-derived exosomes: a novel molecular target for cardioprotection through miR-873-5p/XIAP axis. Cell Death Dis. 2020;11(8):696.
9. Wolska M, Jarosz-Popek J, Junger E, Wicik Z, Porshoor T, Sharif L, Czajka P, Postula M, Mirowska-Guzel D, Czlonkowska A, et al. Long non-coding RNAs as promising therapeutic approach in ischemic stroke: a comprehensive review. Mol Neurobiol. 2020.
10. Zhou Y, Feng D, Gu X, Gao A, Liu Y. The role and clinical significance of long noncoding RNA zinc finger E-box-binding homeobox two antisense RNA 1 in promoting osteosarcoma cancer cell proliferation, inhibiting apoptosis and increasing migration by regulating miR-145. Anticancer Drugs. 2020.
11. Guan X, Xu Y, Zheng J. Long noncoding RNA PCAT6 promotes the development of osteosarcoma by increasing MDM2 expression. Oncol Rep. 2020;44(6):2465–74.
12. Zhang Y, Guo H, Ma L, Chen X, Chen G. Long noncoding RNA LINC00839 promotes the malignant progression of osteosarcoma by competitively binding to microRNA-454-3p and consequently increasing c-Met expression. Cancer Manag Res. 2020;12:8975–87.
13. Shen P, Cheng Y. Long noncoding RNA lncARSR confers resistance to Adriamycin and promotes osteosarcoma progression. Cell Death Dis. 2020;11(5):362.
14. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med. 2013;32(30):5381–97.
15. Yu H, Pan R, Qi Y, Zheng Z, Li J, Li H, Ying J, Xu M, Duan S. LEPR hypomethylation is significantly associated with gastric cancer in males. Exp Mol Pathol. 2020;116:104493.

Yu *et al. BMC Med Genomics*   (2021) 14:124

Page 10 of 10

16. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523.

17. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–13.

18. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45(D1):D362–8.

19. Angulo P, Kaushik G, Subramaniam D, Dandawate P, Neville K, Chastain K, Anant S. Natural compounds targeting major cell signaling pathways: a novel paradigm for osteosarcoma therapy. J Hematol Oncol. 2017;10(1):10.

20. Li Y, Ge F, Wang S. Four genes predict the survival of osteosarcoma patients based on TARGET database. J Bioenergy Biomembr. 2020;52(4):291–9.

21. Zhou K, Xu J, Yin X, Xia J. Long noncoding RNA HAGLROS promotes cell invasion and metastasis by sponging miR-152 and upregulating ROCK1 expression in osteosarcoma. Comput Math Methods Med. 2020;2020:7236245.

22. Wu K, Feng Q, Li L, Xiong Y, Liu S, Liu J, Wu Q. Long-noncoding RNA PCAT6 aggravates osteosarcoma tumourigenesis via the MiR-143-3p/ZEB1 axis. Onco Targets Ther. 2020;13:8705–14.

23. Sun X, Tian C, Zhang H, Han K, Zhou M, Gan Z, Zhu H, Min D. Long noncoding RNA OIP5-AS1 mediates resistance to doxorubicin by regulating miR-137-3p/PTN axis in osteosarcoma. Biomed Pharmacother. 2020;128:110201.

24. Di Mauro S, Scamporrino A, Petta S, Urbano F, Filippello A, Ragusa M, Di Martino MT, Scionti F, Grimaudo S, Pipitone RM, et al. Serum coding and non-coding RNAs as biomarkers of NAFLD and fibrosis severity. Liver Int. 2019;39(9):1742–54.

25. Shen Y, Feng Y, Chen H, Huang L, Wang F, Bai J, Yang Y, Wang J, Zhao W, Jia Y, et al. Focusing on long non-coding RNA dysregulation in newly diagnosed multiple myeloma. Life Sci. 2018;196:133–42.

26. Zheng S, Zhao T, Yuan S, Yang L, Ding J, Cui L, Xu M. Immunodeficiency promotes adaptive alterations of host gut microbiome: an observational metagenomic study in mice. Front Microbiol. 2019;10:2415.

27. Wang X, Jiao X, Tian Y, Zhang J, Zhang Y, Li J, Yang F, Xu M, Yu X, Shanghai Birth Cohort Study. Associations between maternal vitamin D status during three trimesters and cord blood 25(OH)D concentrations in newborns: a prospective Shanghai birth cohort study. Eur J Nutr. 2021. https://doi.org/10.1007/s00394-021-02528-w.

28. Yan X, Zhao X, Li J, He L, Xu M. Effects of early-life malnutrition on neurodevelopment and neuropsychiatric disorders and the potential mechanisms. Prog Neuropsychopharmacol Biol Psychiatry. 2018;83:64–75.

29. Jin G, Xu M, Zou M, Duan S. The processing, gene regulation, biological functions, and clinical relevance of N4-acetylcytidine on RNA: a systematic review. Mol Ther Nucleic Acids. 2020;20:13–24.

30. Kapusta A, Feschotte C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. Trends Genet. 2014;30(10):439–52.

31. Duan H, He H, Hu Q, Lin Y, Cao S, Lan X, Li L, Pang D. Comparison of regulatory networks of E74-like factor 1 and cold-shock domain-containing E1 in breast cancer cell lines using ChIP datasets. Exp Ther Med. 2020;20(6):245.

32. Yang C, Xu W, Gong J, Liu Z, Cui D. Novel somatic alterations underlie Chinese papillary thyroid carcinoma. Cancer Biomark. 2020;27(4):445–60.

33. Zhong D, Cen H. Aberrant promoter methylation profiles and association with survival in patients with hepatocellular carcinoma. Onco Targets Ther. 2017;10:2501–9.

34. Chen J, Zhao X, Cui L, He G, Wang X, Wang F, Duan S, He L, Li Q, Yu X, Zhang F, Xu M. Genetic regulatory subnetworks and key regulating genes in rat hippocampus perturbed by prenatal malnutrition: implications for major brain disorders. Aging (Albany NY). 2020;12(9):8434–58. https://doi.org/10.18632/aging.103150 (**(Epub 2020 May 11)**).

35. Li H, Wang X, Lu X, Zhu H, Li S, Duan S, Zhao X, Zhang F, Alterovitz G, Wang F, Li Q, Tian XL, Xu M. Co-expression network analysis identified hub genes critical to triglyceride and free fatty acid metabolism as key regulators of age-related vascular dysfunction in mice. Aging (Albany NY). 2019;11(18):7620–38. https://doi.org/10.18632/aging.102275 (**(Epub 2019 Sep 12)**).

36. Wu Y, Cao H, Baranova A, Huang H, Li S, Cai L, Rao S, Dai M, Xie M, Dou Y, Hao Q, Zhu L, Zhang X, Yao Y, Zhang F, Xu M, Wang Q. Multi-trait analysis for genome-wide association study of five psychiatric disorders. Transl Psychiatry. 2020;10(1):209. https://doi.org/10.1038/s41398-020-00902-6.

37. Jiang L, Wang K, Lo K, Zhong Y, Yang A, Fang X, Akezhuoli H, Song Z, Chen L, An P, Xu M, Min J, Wang F. Sex-specific association of circulating ferritin level and risk of type 2 diabetes: a dose-response meta-analysis of prospective studies. J Clin Endocrinol Metab. 2019;104(10):4539–51. https://doi.org/10.1210/jc.2019-00495.

38. Liu M, Li F, Yan H, Wang K, Ma Y, Shen L, Xu M, Alzheimer's Disease Neuroimaging Initiative. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. Neuroimage. 2020;208:116459. https://doi.org/10.1016/j.neuroimage.2019.116459 (**(Epub 2019 Dec 16)**).

Publisher's Note

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