Biological analysis of cancer specific microRNAs on function modeling in osteosarcoma

Hao Wang1, Min Tang1, Liping Ou1, Mengyi Hou1, Tianyu Feng1, Yu-E Huang1, Yaqian Jin1, Heng Zhang2 & Guowei Zuo1

Osteosarcoma (OS) is the most common bone tumor characterized with a high risk of amputation and malignant morbidity among teenagers and adolescents. However, relevant pathogenic/biological mechanisms underlying OS-genesis remains to be ambiguous. The aim of this study was to elucidate functional relationship about microRNAs-mRNAs networks and to identify potential molecular markers via a computational method. Gene expression profile (GSE70415) was recruited from Gene Expression Omnibus. 3856 differentially expressed genes and 250 significantly expressed microRNAs were identified by using GCBI. The results of GO and KEGG pathways associated proteomics analysis indicated that extracellular matrix organization, small molecule metabolic process, cell adhesion (GO IDs: 0030198, 0044281, 0007155) and pathways in cancer, PI3K-Akt signaling pathway, metabolic pathways (pathway IDs: 5200, 4151, 1100) were significantly enriched. In addition, CKMT2, miR-93b-5p, miR-29b-3p were found to be positively/negatively correlated with TP53, EGFR, and MMP members mediated OS development, including angiogenesis, migration and invasion. Further visualization of collective effect of 1181 microRNAs-mRNAs pairs and protein-protein interactions was realized by applying with cytoscape. In summary, our work provided a better understanding of non-coding regulatory mechanisms of transcriptomics and unraveled essential molecular biomarkers in osteosarcoma.

Osteosarcoma (OS) is the most frequent primary bone malignancy, characterized with a high potential for lung metastasis and has been the third common cancer-associated threat to adolescents. It most occurs at the extremities of long bones, where osteoblasts transform into mature bone tissue. However, the putative molecular mechanisms underlying OS carcinogenesis have not been deciphered completely and still been a challenge yet. Hitherto, cumulative evidences have demonstrated that a variety of factors including microRNAs (miRNAs), a group of non-coding RNAs, were involved in OS development. The first study on miRNAs expression in OS published by Gao et al.8 identified 182 differentially expressed miRNAs (DEmiRNAs), accelerating revelation that miRNAs may have an obscure but critical impact underlying OS pathogenesis. Recent researches also suggested that miR-1, -409-3p, -379, -665, -489-3p function as sequence-specific tumor suppressors mediating primary OS proliferation, cell death and even distant metastasis.

Alternatively, development of high throughput testing technology (microarray, next-generation sequencing) has successfully made it convenient to acquire large-scale genetic data. Bioinformatics approach uniting biology, mathematics, and computer science has further widely facilitated molecular mechanism explanation and discovery of tumor-correlated diagnostic markers. RNA-sequencing has found that amounts of genes come into discrepancy along the course of bone malignancy transformation. By comparing mRNA expression profiles between OS tissues and cell lines and xenografts, Kuijjer M L et al.14 initially achieved histological subtyping classification (osteoblastic, chondroblastic, fibroblastic) at transcriptome level. In parallel, epigenetic events and RUNX2 interactome were identified to be constitutively activated in OS15.

Nevertheless, targeting networks of miRNAs to mRNAs underlying osteosarcomagenesis have not been systematically interpreted yet. MiRNAs are essential components in biological homeostasis and the current paper

---

1Key Laboratory of Diagnostic Medicine designated by the Chinese Ministry of Education, Department of Laboratory Medicine, Chongqing Medical University, Yuzhong district, Chongqing, 400016, China. 2Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, 60611, USA. Hao Wang and Min Tang contributed equally to this work. Correspondence and requests for materials should be addressed to G.Z. (email: gwzuo@qq.com)
has accomplished miRNAs-involved networks construction and exploitation for essential biomarkers along pipeline of central dogma by means of easy-handling web-based analytical tool GCBI.

Results

DEGs and DEmiRNAs between hMSC and OS cell lines. The sample set GSE70415, which consists of miRNA (GSE70367) and mRNA (GSE70414) expression profiles of five human OS cell lines (MG63, Saos, HOS, NY, Hu09) and a corresponding control (hMSC) was obtained from Gene Expression Omnibus (GEO). Following standard protocol of samples qualification and normalization, raw expression values were summarized and analyzed in a consecutive workflow (seen in Fig. S1) based on GCBI. In total, 3856 (P < 0.01) significant DEGs were identified, of which 1705 over-presented and 2151 showed an attenuated behavior (Fig. 1a). Periostin (POSTN), a canonical osteoblast marker, has not only exhibited a most significant declination among the whole collection, but recent study has also already verified hypothesis that aberrant stimulation of it concerned with bevacizumab induced resistance in the cases of glioma implementing with anti-VEGF-A therapy. Meanwhile 250 (P < 0.01) DEmiRNA picked out from microRNA repertoire comprised by 161 ascent items and 89 down-regulated miRNA episodes (Fig. 1b). Whereas, of some limitation, miR-182-5p and miR-708-5p, existing the highest contradictory deviation (absolute fold change|FC > 100) within current community events both could not be tracked among 81 small sequences in curated Osteosarcoma Database. Along the clarification of microRNA-engaged epigenetic reprogramming, potential connection between both of them and vorinostat, an approved histone deacetylase inhibitor was further validated in 143B and MG63 (data have not been published). In addition, DEGs which were statistically significant complied with cumulate information partially (about 7%) after matching to 911 trustwor-thy entries within the Osteosarcoma Database (seen in Fig. S2). The full tables of DEGs and DEmiRNAs were included in Tables S2 and S3.

Functional enrichment of DEGs and DEmiRNAs between hMSC and OS cell lines. As known, tumorigenesis is featured with a number of biological disorders and cellular events dysregulation, such as angiogenesis, cell adhesion, signaling transduction. Thus, it is absolutely necessary to unravel discrepant biological processes and pathways recruited along the duration of neoplasia. In the enrichment modules, 395 records of GO and 142 KEGG pathways (full tables can be seen in Tables S4 and S5) were verified through employment with Fisher exact testing and FDR. Moreover, we annotated top-ranked 20 GO and KEGG pathways respectively without distinguishing biological process (BP), cell component (CC) and molecular function (MF) (Fig. 2a and b). It is obvious that the top three enriched biological processes contained extracellular matrix organization, small molecule metabolic process, cell adhesion (GO IDs: 0030198, 0044281, 0007155). Whereas, pathways in cancer, PI3K-Akt signaling pathway, metabolic pathways (pathway IDs: 5200, 4151, 1100) were three most significantly concentrated pathways through which oncogenes silencing was switched on or off. Both GO and KEGG pathway enrichment analysis showed a peak distribution of DEGs in metabolic dysfunction. To some extent, this was consistent with previous consensus that tumor events, such as proliferation, metastasis and angiogenesis could be partially attributed to hypermetabolic activity of neoplasm. Besides, MAPK signaling pathway, pathways in cancer, and cell cycle (pathway IDs: 4010, 5200, 4110) acted as leading initiators mediating follow-up aberrant
pathway cascades through assessment of determination coefficient (Fig. 3a). Intensive pathways featured with more than 10 contribution degrees were formatted into Table 1.

**Genes interplay and co-expression networks.** To further explore and clarify realization of message or communication flow from member to member scattered at the crossing pathways, visualization and cluster analysis of hub genes were accomplished using cytoscape 3.4.0. We picked out 698 overlapped genes derived from GO and KEGG pathway analysis and applied them to genesignal (shown in Fig. S4) and co-expression network construction (Fig. 3b). As co-expression graphic illustrated, correlative genes positively or negatively interacted with their neighbors in a non-direction nested manner. According to MCODE20 analyzer, 19 subordinated nodes intimately clustered to creatine kinase, mitochondrial 2 (CKMT2), also known as SMTCK, which was indispensable when maintaining rational energy metabolism. Thus, our colleague later validated hypothesis that CKMT2 might as a key regulating factor participating in osteosarcomagenesis (data have not been published).

**Targets prediction and miRNAs-targets interaction.** MiRNAs, a group of well-known endogenous non-coding RNA, usually act as transcription regulators during gene expression through binding to 3′-untranslated region (3′-UTR) of target mRNAs. It is explicit that diversity of miRNAs resulted from length or alignment of seed region complicates regulatory models. Thus, further understanding net-association between miRNAs and mRNAs is extremely needed. By utilizing GCBI that integrating TargetScan21 and miRanda22 databases, we mined out 250 DEmiRNAs with an up to down ratio at 161/89 (shown in Fig. 1b). Abiding by the base-pairing principle, there were 29227 genes found deposited in the target pools (TargetScan and miRanda). Conversely, 388 were substantially involved in GO enrichment (seen in Fig. S3) and 608 were mingled with DEmiRNAs regardless of exact binding pair bases. To delineate miRNAs-mRNAs axis vividly, we postulated index degrees which changed no less than 10 to be of significance in transcription function in our research and deeply screened impaction networks of selected 40 DEmiRNAs (shown in Table S1). Illustration of connective networks of miRNAs and corresponding targets were realized using cytoscape 3.4.0. In summary, 238 downstream genes were blocked and 181 targets found to be in an activated status (Fig. 4a and b). The results showed that either lower-expressed miR-29b-3p or over-presented miR-93-5p was hub miRNA possessing most significant impact on gene transcription and even protein function implement.

**Protein-protein interaction in OS cell lines.** To study protein-protein interactive association of DEGs mediated by DEmiRs, we screened 35 typical DEmiRs (FC ≥ 10 compared to control) and integrated protein-protein interaction (PPI) network of under-manipulated target mRNAs by means of STRING 10.0. Neither six isolated nodes (has-miR-941, 127-3p, 487b-3p, 34a-3p, 493-3p, 654-3p) without microRNA-mRNA joint nor molecules that absent from function (GO or KEGG) participation was eliminated. Then emerged 43 genes were employed to construct PPI network by using cytoscape 3.4.0 (Fig. 5a). Within shaped model, receptor nodes already have been verified or not, such as FOXO1, BMP, members of COL and ITG families were predicted to interact with members essential for pathway perturbation, among which some classical suppressive factors involved, like TP53, EGFR and MMP2.
**Figure 3.** Co-expression network analysis of osteosarcoma. Significantly coefficient KEGG pathway network (a, left) was visualized with augmented index degrees (circles from cyan to red). Co-expressed DEGs were integrated into networks using bioinformatics methodology (b, right). Positive/negative function among common genes (rectangles, blue) and even tightly clustered elements (purple and green) were displayed with different colors (red and black).

| Pathway ID | Pathway name               | Degree | Pathway feature |
|------------|----------------------------|--------|-----------------|
| 4010       | MAPK signaling pathway     | 34     | down|up            |
| 5200       | Pathways in cancer         | 27     | down|up            |
| 4110       | Cell cycle                 | 24     | down|up            |
| 4210       | Apoptosis                  | 23     | down|up            |
| 4115       | p53 signaling pathway      | 17     | down|up            |
| 4310       | Wnt signaling pathway      | 16     | up|down           |
| 4520       | Adherens junction          | 15     | down|up            |
| 4012       | ErbB signaling pathway     | 14     | down|up            |
| 4350       | TGF-beta signaling pathway | 13     | down|up            |
| 4510       | Focal adhesion             | 12     | down|up            |
| 4060       | Cytokine-cytokine receptor interaction | 12 | down|up            |

**Table 1.** Remarkably correlative pathways in functional enrichment.

**Figure 4.** Regulatory networks of onco-associated miRNAs in osteosarcoma cells. Functional models of up-regulated miRNAs (a, left) and down-regulated miRNAs (b, right) were constructed, respectively. Up/down-regulated miRNAs were exemplified with triangles or circles with red color whereas targeted genes showed with blue squares.
Discussion

In this study, we firstly provided a systematical miRNAs-mRNAs functional model based on expression profiles of OS transcriptome. Distinctive to previous researches focusing on individual command element, we analyzed a large number of molecules and integrated them into a functional network via adopting a bioinformatic approach. This research is not only a promotion in revealing small non-coding RNA disorder hiding in oncogenesis, even chemoresistance, but also indispensable for clinical early-screening and targeted therapy exploitation23, though underveining disturbance mediated by genetic or microenvironment origin remains to be a challenge.

By microarray analysis, we firstly identified 3856 mRNAs and 250 miRNAs which significantly diverged in OS cells. POSTN, mainly involved in osteoblasts adhesion and differentiation24, 25, was found declined remarkably in OS subgroups comparing to normal sets. Nevertheless, what fascinated us was that expression of POSTN had been reported to remain at a high level in OS compared to osteochondroma and high content of POSTN intensely correlated with tumor angiogenesis and poor prognosis in the OS as well as high grade glioma in vivo17, 26, 27. The probable reasons for this discrepancy might be inconsistent of sample type (cell lines versus specimens) and detection approaches (RNA microarray versus immunohistochemistry). Subsequently, the results of functional enrichment analysis demonstrated that metabolic pathway played an important role and a large number of cancer associated pathways were distinguished, including PI3K-Akt and MAPK signaling. There is a reason to believe that chemoresistance is relevant to metabolic abnormality as miR-221, −101, −22, −15528–31 have already been proved to participate in cisplatin and doxorubicin derived chemo-resistance as well as our investigation about SAHA to miR-182-5p and -708-5p in OS cells. Alternatively, activation of the PI3K-Akt pathway suppressed cell longevity through phosphorylation of FOXO members and balancing its activity with MAPK and NF-κB pathway intimately associated with tumors survival1. On the other hand, stimulation of MAPK signaling was confirmed to link with elevated EGFR phosphorylation and MMP-9 levels mediated by lowering miR-143 in OS32. Except those miRNAs-pathways5, 7, 33 verified so far, newly discovered miRNAs expanded OS related miRNAs spectrum notably. Furthermore, modeling of miRNAs-mRNAs networks was achieved using a well-established tool to visualize intricate nodes connections (Fig. 4a and b). Despite not the most altered, miR-29b-3p and miR-93-5p were two core upstream elements targeting transcription proceeding of which miR-29b-3p induced OS depression had been affirmed to be with tumor-specific subcellular localization34 while the most significant miR-182-5p and miR-708-5p displayed relatively moderate and even lower contribution degrees. It seems that efficiency of miRNAs is not simply determined by the level of variation but relied on critical GO and pathways. In summary, 1181 linkages have been established in the current study, which has been a striking acceleration about non-coding unit mediated OS carcinogenesis.

Bioinformatics approach combining GCBI and cytoscape, an innovative pipeline distinguished from troublesome data processing to pattern display, facilitating dimensional molecular interaction and model analysis based on antecedent data and improved algorithm multidisciplinary. It is insufficient that our present paper has just successfully explained the relationship between microRNAs and coding targets, and necessary to further supplement another non-coding factors, including long non-coding RNAs, circle RNAs mediated competitive mechanism. There is no doubt that combinational strategies through employing identification of group effect of non-coding RNAs-miRNAs-proteins even small inhibitors and drugs would be potent approaches and might bring a breakthrough.
Methods

Data source. In the present paper, gene expression profile (accession: GSE70415) deposited in the Gene Expression Omnibus (GEO)38 (https://www.ncbi.nlm.nih.gov/geo/) was utilized. GSE70415 comprised a total of five OS samples (MG63, Saos, HOS, NY, Hu09) and a normal control (hMSC). Subset expression profiles of mRNAs (GSE70414) and miRNAs (GSE70367) were detected using Affymetrix Human Genome U133 Plus 2.0 Array and Affymetrix Multispecies miRNA-3 Array (Inc. Santa Clara, CA, USA) based on platform GPL 570 and 16384. Totally, 54675 genes and 25533 miRNAs were mined, respectively.

Differentially expressed genes and miRNAs-mRNAs analysis. To identify DEGs and DEmiRNAs between OS and hMSC cell lines, a web-based online tool GCBI (www.gcbi.com.cn/gcbi/html/index) was utilized. Entries qualification and calibration were then achieved by taking standard median Polish algorithm16. Only probe signals with p-values < 0.01, false discovery rate (FDR) < 0.01 and absolute value of fold change (FC) > 2 were considered to be statistically differential. Genesignal and co-expression network were further constructed based on contribution degrees according GCBI protocol (http://college.gcbi.com.cn/helpme).

Enrichment analysis and networks construction. For visualization, cytoscape 3.4.036 (http://www.cytoscape.org/), an open source platform, was utilized to portray the relationship among target molecules. Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis18 for DEGs was performed with GCBI mentioned above. DEGs and DEmiRNAs given previously were selected out to construct networks. Molecular Complex Detection (MCODE)20, based on vertex weighting by local neighborhood density, was used to group gene modules.

Databases. To identify putative regulatory correlation within our work, some essential databases were consolidated to explore the relevance of non-coding transcriptome to proteomics. STRING10.039 (http://string-db.org/) was implemented to supervise protein–protein interactions with. MiRNAs which associated with oncogenesis were further filtered according to osteosarcoma9 (http://osteosarcoma-db.uni-muenster.de/) and oncomiRDB8 (http://bioinfo.au.tsinghua.edu.cn/member/jgu/oncomirdb/).

References

1. He, J. P. et al. Review of the molecular pathogenesis of osteosarcoma. Asian Pacific Journal of Cancer Prevention Apjc 15, 5967 (2014).
2. Carvalho, I. N. S. R. D. & Freitas, R. M. D. Translating microRNAs into biomarkers: What is new for pediatric cancer? Medical Oncology 33, 49 (2016).
3. Kushlinskii, N. E., Fridman, M. V. & Braga, E. A. Molecular Mechanisms and microRNAs in Osteosarcoma Pathogenesis. Biochemistry 81, 315 (2016).
4. Sampson, V. B., Yoo, S., Kumar, A., Vetter, N. S. & Kolb, E. A. MicroRNAs and Potential Targets in Osteosarcoma: Review. Frontiers in Pediatrics 3, 69 (2015).
5. Poos, K. et al. Structuring osteosarcoma knowledge: an osteosarcoma–gene association database based on literature mining and manual annotation. Database the Journal of Biological Databases & Curations 2014, 2159–2160 (2014).
6. Nandas, H. M. et al. Modulation of the osteosarcoma expression phenotype by microRNAs. PloS one 7, 950–953 (2012).
7. Server, A. L., Phalak, R., Thayyil, V. & Subramanian, S. S. MED: microRNA expression database. Laboratory Investigation 90, 753–761 (2010).
8. Gao, J. et al. [Cloning and identification of microRNA from human osteosarcoma cell line SOSP-9607]. Chinese Journal of Cancer Research 26, 561–565 (2007).
9. Niu, J., Sun, Y., Guo, Q., Niu, D. & Liu, B. miR-1 Inhibits Cell Growth, Migration, and Invasion by Targeting VEGFA in Osteosarcoma Cells. Disease Markers 2016, 706896 (2016).
10. Wu, S. et al. MicroRNA-409-3p inhibits osteosarcoma cell migration and invasion by targeting catenin-β1. Gene 584, 83–89 (2016).
11. Z., L., J., S., M., T., C. & W., K. MicroRNA-379 suppresses osteosarcoma progression by targeting PDK1. Journal of Cellular and Molecular Medicine (2016).
12. Liu, Q., Yang, G. & Qian, Y. Loss of MicroRNA-489-3p promotes osteosarcoma metastasis by activating PAX3-MET pathway. Molecular carcinogenesis (2016).
13. Yang, Y. Identification of differentially expressed genes in the development of osteosarcoma using RNA-seq. Oncotarget (2016).
14. Kuijjer, M. L. et al. mRNA expression profiles of primary high-grade central osteosarcoma are preserved in cell lines and xenografts. BMC medical genomics 4, 1–12 (2011).
15. Sadikovic, B. et al. Identification of interactive networks of gene expression associated with osteosarcoma oncogenesis by integrated molecular profiling. Human molecular genetics 18, 1962–1975, doi:10.1093/hmg/ddp117 (2009).
16. Kapur, K., Xing, Y., Ouyang, Z. & Wong, W. H. Exon arrays provide accurate assessments of gene expression. Genome biology 8, R82, doi:10.1186/gb-2007-8-5-r82 (2007).
17. Park, S. Y., Piao, Y., Jeong, K. J., Dong, J. & de Groot, J. F. Periostin (POSTN) regulates tumor resistance to antiangiogenic therapy in glioma models. Molecular Cancer Therapeutics 15 (2016).
18. Benjamin, Y. & Hochberg, Y. Controlling The False Discovery Rate - A Practical And Powerful Approach To Multiple Testing. Journal of the Royal Statistical Society 57, 289–300 (1995).
19. Ren, L. et al. Metabolomics uncovers a link between inositol metabolism and osteosarcoma metastasis. Oncotarget, doi:10.18632/oncotarget.15872 (2017).
20. Bader, G. D. & Hogue, C. W. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 4, 2 (2003).
21. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. eLife 4, doi:10.7554/eLife.05005 (2015).
22. Betel, D., Wilson, M., Gabow, A., Marks, D. S. & Sander, C. The microRNA.org resource: targets and expression. Nucleic Acids Res 36, D149–153, doi:10.1093/nar/gkn995 (2008).
23. Ram Kumar, R. M., Boro, A. & Fuchs, B. Involvement and Clinical Aspects of MicroRNA in Osteosarcoma. International Journal of Molecular Sciences 17 (2016).
24. Chuanyu, S., Yuqing, Z., Chong, X., Guowei, X. & Xiaojun, Z. Periostin promotes migration and invasion of renal cell carcinoma through the integrin/local adhesion kinase-c-Jun N-terminal kinase pathway. Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine 39, 1010428317694549, doi:10.1177/1010428317694549 (2017).
25. Wu, Z. et al. Periostin Promotes Migration, Proliferation, and Differentiation of Human Periodontal Ligament Mesenchymal Stem Cells. Connective tissue research, doi:10.1080/03008207.2017.1306060 (2017).

26. Hu, F. et al. High-level expression of periostin is significantly correlated with tumour angiogenesis and poor prognosis in osteosarcoma. International Journal of Experimental Pathology 97, 86–92 (2016).

27. Hu, F., Wang, W., Zhou, H. C., & Shang, X. F. High expression of periostin is dramatically associated with metastatic potential and poor prognosis of patients with osteosarcoma. World Journal of Surgical Oncology 12, 287–287 (2014).

28. Zhao, G. et al. MicroRNA-221 Induces Cell Survival and Caspase-Resistance through PI3K/Akt Pathway in Human Osteosarcoma. e53906 (2013).

29. Chang, Z., Huo, L., Li, K., Wu, Y. & Hu, Z. Blocked Autophagy by miR-101 Enhances Osteosarcoma Cell Chemosensitivity In Vitro. TheScientificWorldJournal 2014, 794756–794756 (2014).

30. Li, X., Wang, S., Chen, Y., Liu, G. & Yang, X. miR-22 targets the 3′ UTR of HMGB1 and inhibits the HMGB1-associated autophagy in osteosarcoma cells during chemotherapy. Tumor Biology 35, 6021–6028 (2014).

31. Chen, L., Jiang, K., Jiang, H. & Wei, P. miR-155 mediates drug resistance in osteosarcoma cells via inducing autophagy. Experimental & Therapeutic Medicine 8, 527–532 (2014).

32. Wang, Q., Cai, J., Wang, J., Xiong, C. & Zhao, J. MiR-143 inhibits EGFR-signaling-dependent osteosarcoma invasion. Tumor Biology 35, 12743–12748 (2014).

33. Wang, D., Gu, J., Wang, T. & Ding, Z. OncomiRDB: a database for the experimentally verified oncogenic and tumor-suppressive microRNAs. Bioinformatics 30, 2237–2238 (2014).

34. Jones, K. B. et al. miRNA signatures associate with pathogenesis and progression of osteosarcoma. Cancer Res 72, 1865–1877, doi:10.1158/0008-5472.CAN-11-2663 (2012).

35. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research 30, 207–210 (2002).

36. Shannon, P. et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction. Networks 13, 2498–2504 (2003).

37. Szklarczyk, D. et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43, D447–452, doi:10.1093/nar/gku1003 (2015).

Acknowledgements

This study was supported by the National Natural Science Foundation of China (81102035). The authors wish to thank the technical assistance of the online analysis platform GCBI website (http://www.gcbi.com.cn).

Author Contributions

G.Z. conceived and instructed the work. H.W. and M. performed the analysis and constructed the model. L.O., M.H., T.F., Y.H. and Y.J. checked the associated database and classified raw data. H.W., M.T., H.Z. and G.Z. wrote and revised the manuscript. All of the authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-05819-7

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

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017