



Research Paper

Prediction of optimal gene functions for osteosarcoma using gene ontology and microarray profiles



Xinrang Chen

Pediatric Surgery, First Affiliated Hospital of Zhengzhou University, NO. 1 Jianshe East Road, Zhengzhou, Henan 450052, People's Republic of China

ARTICLE INFO

Keywords:

Osteosarcoma
Spearman correlation coefficient
Gene ontology
Area under the curve
Differential co-expression network
Guilt by association

ABSTRACT

In the current study, we planned to predict the optimal gene functions for osteosarcoma (OS) by integrating network-based method with guilt by association (GBA) principle (called as network-based gene function inference approach) based on gene ontology (GO) data and gene expression profile. To begin with, differentially expressed genes (DEGs) were extracted using linear models for microarray data (LIMMA) package. Then, construction of differential co-expression network (DCN) relying on DEGs was implemented, and sub-DCN was identified using Spearman correlation coefficient (SCC). Subsequently, GO annotations for OS were collected according to known confirmed database and DEGs. Ultimately, gene functions were predicted by means of GBA principle based on the area under the curve (AUC) for GO terms, and we determined GO terms with AUC > 0.7 as the optimal gene functions for OS. Totally, 123 DEGs and 137 GO terms were obtained for further analysis. A DCN was constructed, which included 123 DEGs and 7503 interactions. A total of 105 GO terms were identified when the threshold was set as AUC > 0.5, which had a good classification performance. Among these 105 GO terms, 2 functions had the AUC > 0.7 and were determined as the optimal gene functions including angiogenesis (AUC = 0.767) and regulation of immune system process (AUC = 0.710). These gene functions appear to have potential for early detection and clinical treatment of OS in the future.

1. Introduction

Osteosarcoma (OS) is the most commonly diagnosed histological form of primary bone tumor with high morbidity, and it is mainly prevalent in teenagers and young people [1]. Pain is the most common early symptom of OS and can cause fracture of the affected bone. Currently, multiple therapeutic strategies for OS, for example, surgical resection, chemotherapy and radiotherapy, have significantly improved the prognosis of patients with OS [2]. However, the overall survival rates rarely exceed 60–65% [3]. Moreover, a significant portion of OS patients develop metastasis even after curative resection of the primary tumor. There is still a long way to go for the management of OS [4]. Therefore, with an attempt to continue to make progress in the diagnosis and management of OS, identification of sensitive and specific minimally invasive signatures is one of the most important challenges.

Genetic aberration has been demonstrated as an important factor that may play significant roles in OS pathogenesis. For instance, Zhu and colleagues [5] have reported that *SOX9* is over-expressed in OS tissues compared with controls. Moreover, it is indicated that *FOXMI* is up-regulated, which suggests *FOXMI* might be an valuable bio-signature for OS [6]. Nevertheless, these genes do not treat the OS efficiently

or selectively, because molecules frequently don't work individually, yet co-operate with the other genes. Additionally, genetic factors can disturb the protein levels, thereby in turn perturb the molecule interactions. Network is characterized by the complicated interactions and the complex interwoven relationships that control cellular functions [7]. Hence, understanding the networks will be beneficial to provide new insights to explore the molecular pathogenesis of OS. The concept of differential co-expression network (DCN) has been employed to the studies of OS, due to statistical confidence of single connections, overlap with protein interaction, and mathematical convenience [8]. In addition, improving our knowledge of gene function in uncharacterized genes is a major task [9]. Remarkably, gene interactions can be applied to deduce functional relationships based on a principle known as “guilt by association” (GBA) [10]. GBA has been indicated to predict gene function in various types of biological networks, for example, gene co-expression network [11].

Thus, in our work, we integrated network-based method with GBA principle (called as network-based gene function inference approach) to further identify the optimal gene functions for OS using gene ontology (GO) data and gene expression profile. To begin with, differentially expressed genes (DEGs) were extracted using linear models for microarray data (LIMMA) package. Then, construction of DCN based on DEGs

E-mail address: chenxr689@163.com.

<http://dx.doi.org/10.1016/j.jbo.2017.04.003>

Received 15 March 2017; Received in revised form 5 April 2017; Accepted 6 April 2017
Available online 08 April 2017

2212-1374/ © 2017 The Author. Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

was implemented, and sub-DCN was identified using Spearman correlation coefficient (SCC). Subsequently, GO annotations for OS were collected according to known confirmed database and DEGs. Ultimately, gene functions were predicted by means of GBA principle based on the area under the curve (AUC) for GO terms, and we determined GO terms with AUC > 0.7 as the optimal gene functions for OS. These gene functions appear to have potential for early detection and clinical treatment of OS in the future.

2. Materials and methods

2.1. Gene expression profile and data pre-treatment

OS-related microarray data (accession number E-GEOD-36001) [12] were downloaded from ArrayExpress database based on the platform of A-MEXP-930-Illumina Human-6 v2 Expression BeadChip. There were 19 OS samples and 6 normal samples in E-GEOD-36001. Prior to analysis, we firstly pre-processed the microarray profile of E-GEOD-36001. Specifically, robust multi-array average (RMA) was used to perform background correction [13]. Quartile algorithm was utilized to implement quartile normalization [14], following by perfect match (PM)/mismatch (MM) correction using microarray suite (MAS) 5.0 package [15]. Ultimately, the data on probe levels were converted into gene symbols relying on annotate package [16]. Overall, 19,027 genes were reserved for further exploitation.

2.2. Identification of DEGs

The LIMMA package [17] in R language and an empirical Bayes framework were applied in our analysis to achieve DEGs between OS and normal samples. A *t*-test was conducted, and the multiple test was applied to correct the raw P values using the Benjamini & Hochberg [18] method based on false discovery rate (FDR). DEGs were extracted on the basis of FDR < 0.05 and |log fold change (FC)| ≥ 2.

2.3. Generation of DCN

Cytoscape (<http://cytoscape.org/>) is an open source software which integrates bio-molecular interactions with high-throughput expression data as well as other molecular states into a unified conceptual network [19]. Therefore, we inputted DEGs into the Cytoscape software to visualize the DCN. Furthermore, with the goal of assessing the co-expressed strength of each interaction in the DCN, SCC was used in our study. As we all know, SCC is used to measure the co-expression probability of two variables by assessing the strength of association of two co-expressed variables and it ranges from -1 to 1 inclusive [20,21]. The SCC absolute value of one interaction was determined as the weight value of the corresponding edge, and the higher the weight value was, the more relevant the interaction was related to the disease. Thus, we selected the edges with weight values higher than 0.8 to construct the sub-DCN. In order to display the sub-DCN more vividly, Cytoscape software was applied.

2.4. GO annotation for DEGs

GO annotation has been broadly utilized as functional enrichment studies for large-scale genes [22]. In our study, human GO annotations comprised of 19,003 GO terms covering 18,402 genes were retrieved from the GO consortium (<http://geneontology.org/>). In an attempt to obtain stable performance, we filtered for the GO annotations on gene size such that each remaining GO term had associated genes > 20, and a total of 1313 GO terms were left to be used in our analysis. Next, DEGs identified above were mapped to the 1313 GO annotations. Finally, the GO slim set was obtained, consisting of 123 DEGs and 137 GO terms.

2.5. Identifying gene functions based on “GBA” prediction

Gene networks have been widely used in gene function prediction algorithms, many based on complex extensions of the “GBA” principle. As demonstrated here, GBA prediction approach was utilized to predict significant gene functions in OS progression. For GBA method, we used three-fold cross-validation to extract a ranked list scoring genes in DCN as to how they belonged in the known GO terms. The sum of co-expression values between the training set (co-expression) and the candidate gene was divided by the sum of co-expression values between the genes outside the training set and the candidate gene to analyze degree of candidacy. In detail, for each gene *i* in the DCN, all other neighbored genes of gene *i* were mapped to each GO category *K*, and the multifunctionality (MF) score for each gene *i* within the *K*-GO term based on the following equation:

$$MF(\text{gene}_i) = \sum_{i \in \text{gene}_A} \frac{1}{N_{inK}} * N_{outK}$$

In this formula, N_{in_i} stood for the number of genes in GO term *i*, and N_{out_i} denoted the count of genes outside GO category *i*.

Then, based on support vector machine (SVM), AUC for each GO group *K* was calculated, and the mean AUC across all GO terms was determined. Thus, the AUC scores were ordered from the highest to the lowest, the ranks of GO terms were ordered oppositely. The AUC of 0.5 stands for classification at chance levels, while the AUC of 1.0 denotes a perfect classification. In the literature about the gene function prediction, the AUC greater than 0.7 are regarded good [23]. In our study, GO terms with AUC > 0.7 were identified and regarded as the optimal gene functions.

3. Results

3.1. DCN construction

Before DCN generation, we firstly identified DEGs between OS and normal samples using *t*-test. Based on FDR < 0.05 and |log FC| ≥ 2, a total of 123 DEGs were identified. The top 20 DEGs were shown in Table 1. The most significant DEGs were *HOXB7*, *RHPN2*, *SRGN*, *FOXF2*, and *PLVAP*.

In order to further explore the biological activities of DEGs, we constructed a DCN using the above-identified 123 DEGs. In this DCN, there were 123 nodes and 7503 interactions. Under the DCN, node degree was more than just an statistic about a gene, and significantly,

Table 1
List of the top 20 differentially expressed genes (DEGs).

Genes	log (fold change)	False discovery rate (FDR)
HOXB7	2.085898	1.33E-07
RHPN2	2.147406	6.32E-07
SRGN	5.174341	3.34E-06
FOXF2	2.318869	4.51E-06
PLVAP	3.404251	1.98E-05
COX7A1	2.988735	5.41E-05
APOE	3.399852	6.21E-05
SPINT2	2.301193	6.65E-05
LXN	2.523381	8.93E-05
TNFRSF1B	2.427194	1.10E-04
VAMP8	3.494650	1.61E-04
CBS	2.868625	1.76E-04
HCLS1	2.686705	1.81E-04
PHGDH	2.362045	2.15E-04
GIMAP7	2.258177	2.74E-04
C1QC	2.885076	2.76E-04
HBB	5.191561	2.86E-04
C1QA	3.521570	2.93E-04
TYROBP	3.511873	3.17E-04
C1QB	3.184835	3.17E-04

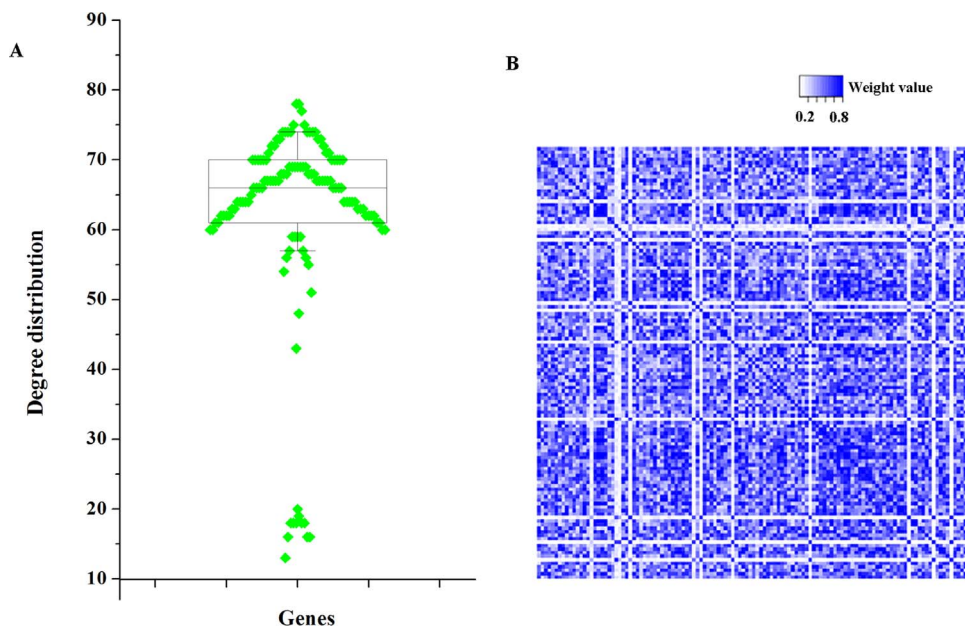


Fig. 1. Differentially co-expressed network (DCN) construction for osteosarcoma (OS) based on differentially expressed genes (DEGs). A. Degree distribution of genes in the DCN. B. Weights distribution of edges in the DCN. Heatmap clarified weight distribution for each interaction.

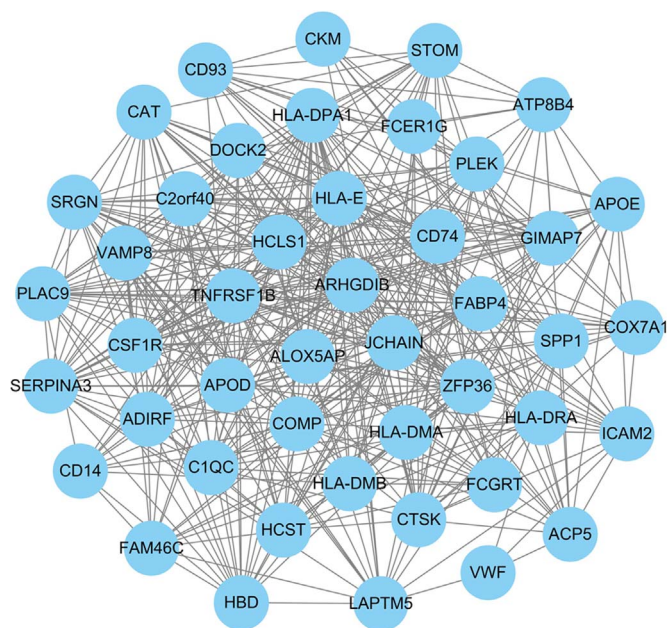


Fig. 2. Sub-DCN using the cut-off threshold of weight value > 0.8. In the sub-DCN, there were 46 nodes and 438 interactions.

degree could explain the structures of the network. Then, topological degree for each node was computed through summing up the nodes it linked directly, and the degree distribution for each node was shown in Fig. 1A. We found that the ranges of degrees for a large number of DEGs (about 79%) was from 59 to 74, and genes (*TNFRSF1B* and *HLA-E*) had the highest degrees of 78. In addition to degree, the interacted strength was an index to estimate the interactions in the DCN. Thus, we used SCC to assign a weight value to each edge, and the specific information were demonstrated in Fig. 1B where a square stood for an edge in the DCN. The deeper the color was, the larger the weight value was. Significantly, among the interactions, there was a good liner correlation which indicated that the DCN had a good network scale property. Then, we selected the interactions with weight values > 0.8 to establish the sub-DCN which was revealed in Fig. 2. In the sub-DCN, there were 46 nodes and 438 interactions. The degree distribution was shown in

Table 2
Degree distribution of 46 genes in the sub-differential co-expression network (DCN).

Genes	Degree	Genes	Degree
TNFRSF1B	33	HBD	20
ZFP36	32	STOM	19
HLA-E	31	DOCK2	18
ARHGDI1B	28	HLA-DMA	17
HLA-DPA1	28	APOE	17
VAMP8	26	LAPTM5	17
FABP4	24	ICAM2	17
GIMAP7	24	FCGRT	16
CD74	24	ACP5	15
PLEK	23	APOD	15
COMP	23	COX7A1	15
HCLS1	23	HLA-DMB	15
PLAC9	23	CD93	14
JCHAIN	23	ATP8B4	14
SERPINA3	23	FAM46C	14
C2orf40	22	CD14	14
SRGN	22	SPP1	13
CTSK	21	HLA-DRA	12
CSF1R	21	FCER1G	12
CAT	21	ALOX5AP	12
ADIRF	21	CKM	8
HCST	21	VWF	4
C1QC	21		

Table 2. Interestingly, in this sub-DCN, *TNFRSF1B* also had the highest degree (degree = 33).

3.2. Identifying gene functions using GBA prediction

In our study, the MF score was first calculated for each gene in a GO term, which influenced the counting membership in a given GO term by how much the gene contributed to the corresponding GO function. The higher the MF score of a gene was, the greater extent to which it should be a good candidate for owing any biological process. Consequently, a single ranked list of genes that best captured candidacy across all functions was equivalent to a list of genes sorted by MF score. Intuitively, if one was forced to choose a single ranking, the gene with the most GO annotations could be predicted as being in all GO categories. This was because if one gene was enriched in 100 GO terms

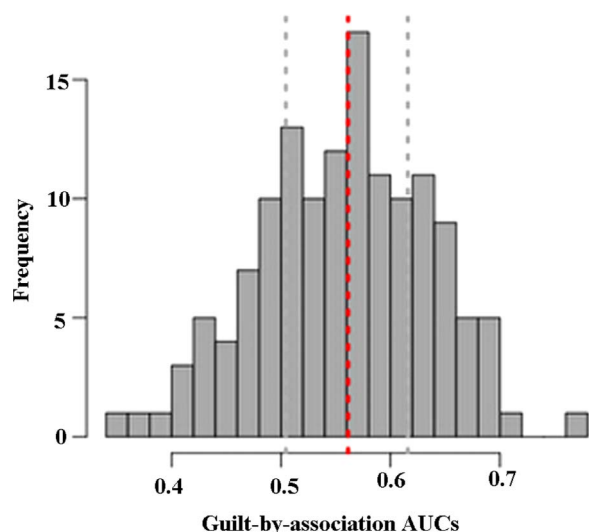


Fig. 3. Gene function prediction performance using guilt by association (GBA). The histogram of AUCs across all GO terms which can be obtained using a single list constructed from number of coexpression partners.

(high MF score), and another was in only one (low MF score), by placing the former gene ahead of the latter gene in a fixed ranking, we made a correct prediction more often across all GO categories. Thus, with the goal of classifying OS patients and normal controls, 3-fold cross-validation was utilized to compute the AUC for GO functions based on MF scores.

The AUC distribution for GO terms was shown in Fig. 3. The AUC for the majority of functional terms ranged from 0.4 to 0.7, especially from 0.5 to 0.65. AUC was used as a predictor to select significant GO terms, so, a total of 105 GO terms were identified based on $AUC > 0.5$. Moreover, among these 105 GO terms, 2 GO terms had the $AUC > 0.7$ and were determined as the optimal gene functions including angiogenesis ($AUC = 0.767$) and regulation of immune system process ($AUC = 0.710$).

4. Discussion

Compared to gene focused researches, gene set and gene signature focused functional investigations appear much rewarding in understanding functional insights [24], and integrative functional genomic analysis of tumors promotes the understanding of the molecular mechanisms of cancers [25]. Network modeling on the basis of co-expression pattern analysis has been widely applied in a variety of cancers to understand the biology processes of cancers, and to obtain clinical insights [26]. Recently, a large amount of techniques have been created to extend GBA to indirect connections to predict gene functions [27,28]. Gene interactions can be applied to deduce functional relationships based on a principle known as GBA [10]. The GBA principle forms the basis for most gene function prediction methods, which typically use relational information (e.g. interactions) to predict new gene membership in gene function categories [29]. However, combination of gene function prediction and network analysis are sparse. Generally speaking, network based-GBA analysis method can make exhaustively examining issues faster and easier than the simple GBA approach. Thus, in this work, we combined GBA principle with DCN-based analysis to further explore both direct and indirect optimal gene functions for OS on the basis of GO information as well as gene expression data. Finally, a total of 105 GO terms were identified based on $AUC > 0.5$, which had a good classification ability. Moreover, 2 out of 105 GO terms had the $AUC > 0.7$ and were determined as the optimal gene functions including angiogenesis and regulation of immune system process.

The ability of tumors to progress to more malignant phenotypes depends on tumor microenvironment. Blood capillaries, made up of

endothelial cells, are usually static under physiological conditions but are induced to proliferate by direct exposure to angiogenic factors, for instance, vascular endothelial growth factor (VEGF) [30]. The proliferation and migration of endothelial cells play an important role in tumor angiogenesis [31]. As we all know, angiogenesis, development of new blood vessels from preexisting vascular bed, plays important roles during tumor growth and metastasis [32]. Significantly, the concept that malignant neoplasm onset, growth, and invasion rely on angiogenesis is broadly recognized and accepted [33,34]. Theoretically, cancer cells, like normal cells, also need the delivery of oxygen as well as nutrients through blood vessels to survive and grow. In situ carcinoma, there seems to be a prolonged dormant period during which the cancer is not angiogenic, and is restrained in growth [35]. When sufficient cancer cells have transformed to the angiogenic phenotype from the quiescent phenotype, neovascularization may initiate, and thus rapid tumor growth can proceed. Through the literatures, we found that gene function of ‘angiogenesis’ has already been established to be important for OS. For example, previous studies have implicated that high recurrence rate and metastatic potential of OS are connected with high levels of angiogenesis [36,37]. Moreover, Wang et al. [38] have used differential co-expression network to find that angiogenesis is closely related to OS progression and metastasis. Most importantly, suppression of cancer angiogenesis has been regarded as an attractive strategy for tumor therapy [39]. Thus, anti-angiogenic therapy might be a consideration for OS patients, with the majority of OS showing high vascularization [40].

During the tumor occurrence and metastasis, cancer cells can modulate the immune system of the host to find strategies that enable them to survive from the immune surveillance [41]. Cancer immunoeediting is one of the two major strategies to get away from immune surveillance [42]. Innate and adaptive immunity appear to contribute to cancer immunoeediting [43]. Previously, lymphocytes and $IFN\gamma$ have indicated to prevent tumor immunoeediting, thereby preventing the selection of less immunogenic tumor cells [44]. Indeed, immune cells (for example, macrophages cells) release soluble agents like chemokines and cytokines promoting the migration and infiltration of leukocytes that exert important functions in tumor development [45]. OS is frequently infiltrated by immune cells including macrophages and T cells [46]. Moreover, macrophage migration inhibitory factor is a proinflammatory cytokine which exerts an crucial function in the immune system [47]. Macrophage migration inhibitory factor contributes to cell proliferation, survival, and tumor-related angiogenesis [48,49]. Moreover, Han et al. [50] have suggested that macrophage migration inhibitory factor can serve as a prognostic marker and a potential therapeutic target for OS. Interestingly, a link between p53 and IFN system was recently discovered in regulating tumor suppression and immunity [51]. Abnormalities of p53 gene in OS occur with a high incidences approaching 50% of all cases [52]. Currently, growing evidence implicates that the immune system is a fundamental player in cancer and a key determinant of prognosis and response to therapy [53,54]. While, understanding the crosstalk between OS cells and the immune system, and how they drive tumorigenesis is still in its infancy. Hence, it is urgently needed to develop novel drugs that would potentiate the immune system in OS patients to act against this disease by means of immunomodulatory methods.

The current study had several limitations. First, there were limited clinical samples which might lead to biased estimates. Second, our analysis was conducted using bioinformatics methods but the conclusions have not been verified by means of any lab technique. Last but not least, we did not compare the obtained findings using other approaches. Despite these limitations, our findings had important implications for the molecular mechanisms underlying OS, yet further experimental study is still needed to validate our study.

In a nutshell, using network-based method with GBA based on GO and microarray profile of OS, we extracted 2 optimal gene functions during OS progression including angiogenesis and regulation of im-

immune system process which might contribute to the successful identification of therapeutic targets for OS. Further studies might shed novel insights into the role of the gene functions in the pathophysiology of OS.

Conflicts of Interest

We declare that there are no conflicts of interest.

References

- [1] K.B. Jones, Z. Salah, D.M. Sara, M. Galasso, E. Gaudio, G.J. Nuovo, et al., MicroRNA signatures associate with pathogenesis and progression of osteosarcoma, *Cancer Res.* 72 (2012) 1865–1877.
- [2] J. Yang, W. Zhang, New molecular insights into osteosarcoma targeted therapy, *Curr. Opin. Oncol.* 25 (2013) 398–406.
- [3] C. Folio, M. Zalacain, C. Zanduetta, C. Ormazábal, L. Sierrasesúmaga, J.M. San, et al., Cortactin (CTTN) overexpression in osteosarcoma correlates with advanced stage and reduced survival, *Cancer Biomark.* 10 (2011) 35.
- [4] M. Hameed, H. Dorfman, Primary Malignant Bone Tumors—Recent Developments in Seminars in Diagnostic Pathology, Elsevier, Amsterdam, Holland, 2011.
- [5] H. Zhu, T. Jie, M. Tang, H. Cai, Upregulation of SOX9 in osteosarcoma and its association with tumor progression and patients' prognosis, *Diagn. Pathol.* 8 (2013) 183–183.
- [6] C.L. Fan, J. Jiang, H.C. Liu, D. Yang, Forkhead box protein M1 predicts outcome in human osteosarcoma, *Int. J. Clin. Exp. Med.* 8 (2015) 15563–15568.
- [7] S.-Y. Sun, Z.-P. Liu, T. Zeng, Y. Wang, L. Chen, Spatio-temporal analysis of type 2 diabetes mellitus based on differential expression networks, *Sci. Rep.* (2013) 3.
- [8] S. Ma, S. Shah, H.J. Bohnert, M. Snyder, S.P. Dineshkumar, Incorporating motif analysis into gene co-expression networks reveals novel modular expression pattern and new signaling pathways, *PLoS Genet.* 9 (2013) 2184–2196.
- [9] D. Eisenberg, E.M. Marcotte, I. Xenarios, T.O. Yeates, Protein function in the post-genomic era, *Nature* 405 (2000) 823–826.
- [10] S. Oliver, Proteomics: guilt-by-association goes global, *Nature* 403 (2000) 601–603.
- [11] L. Peña-Castillo, M. Tasan, C.L. Myers, H. Lee, T. Joshi, C. Zhang, et al., A critical assessment of Mus musculus gene function prediction using integrated genomic evidence, *Genome Biol.* 9 (Suppl. 1) (2008) S2.
- [12] S.H. Kresse, H. Rydbeck, M. Skárn, H.M. Namlós, A.H. Barraganpolania, A.M. Cletonjansen, et al., Integrative analysis reveals relationships of genetic and epigenetic alterations in osteosarcoma, *PLoS One* 7 (2012) e48262–e48262.
- [13] R.A. Irizarry, B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, et al., Exploration, normalization, and summaries of high density oligonucleotide array probe level data, *Biostatistics* 4 (2003) 249–264.
- [14] B.M. Bolstad, R.A. Irizarry, M. strand, T.P. Speed, A comparison of normalization methods for high density oligonucleotide array data based on variance and bias, *Bioinformatics* 19 (2003) 185–193.
- [15] S.D. Pepper, E.K. Saunders, L.E. Edwards, C.L. Wilson, C.J. Miller, The utility of MAS5 expression summary and detection call algorithms, *BMC Bioinforma.* 8 (2007) 1.
- [16] L.J. Zhu, C. Gazin, N.D. Lawson, H. Pagès, S.M. Lin, D.S. Lapointe, et al., ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data, *BMC Bioinforma.* 11 (2010) 1–10.
- [17] G.K. Smyth, Limma: Linear Models for Microarray Data, in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Springer, Berlin, Germany, 2005, pp. 397–420.
- [18] Y. Benjamini, D. Drai, G. Elmer, N. Kafkafi, I. Golani, Controlling the false discovery rate in behavior genetics research, *Behav. Brain Res.* 125 (2001) 279–284.
- [19] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res.* 13 (2003) 2498–2504.
- [20] J. Hauke, T. Kossowski, Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data, *Quaest. Geogr.* 30 (2011) 87–93.
- [21] M.M. Mukaka, Statistics corner: a guide to appropriate use of correlation coefficient in medical research, *Malawi Med. J.: J. Med. Assoc. Malawi* 24 (2012) 69–71.
- [22] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, et al., Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, *Nat. Genet.* 25 (2000) 25–29.
- [23] J. Gillis, P. Pavlidis, The role of indirect connections in gene networks in predicting function, *Bioinformatics* 27 (2011) 1860–1866.
- [24] I. Ponomarev, S. Wang, L. Zhang, R.A. Harris, R.D. Mayfield, Gene co-expression networks in human brain identify epigenetic modifications in alcohol dependence, *J. Neurosci. Off. J. Soc. Neurosci.* 32 (2012) 1884–1897.
- [25] M. Gerlinger, A.J. Rowan, S. Horswell, J. Larkin, D. Endesfelder, E. Gronroos, et al., Intratumor heterogeneity and branched evolution revealed by multiregion sequencing, *N. Engl. J. Med.* 366 (2012) 883–892.
- [26] T.N. Doig, D.A. Hume, T. Theocharidis, J.R. Goodlad, C.D. Gregory, T.C. Freeman, Coexpression analysis of large cancer datasets provides insight into the cellular phenotypes of the tumour microenvironment, *BMC Genom.* 14 (2013) 469.
- [27] H.N. Chua, W.K. Sung, L. Wong, Exploiting indirect neighbours and topological weight to predict protein function from protein-protein interactions, *Bioinformatics* 22 (2008) 1623–1630.
- [28] A.M. Yip, S. Horvath, Gene network interconnectedness and the generalized topological overlap measure, *BMC Bioinforma.* 8 (2007) 22.
- [29] S. Mostafavi, D. Ray, D. Warde-Farley, C. Grouios, Q. Morris, GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function, *Genome Biol.* 9 (Suppl. 1) (2008) S4.
- [30] F. Bussolino, A. Mantovani, G. Persico, Molecular mechanisms of blood vessel formation, *Trends Biochem. Sci.* 22 (1997) 251–256.
- [31] D.W. Leung, G. Cachianes, W.J. Kuang, D.V. Goeddel, N. Ferrara, Vascular endothelial growth factor is a secreted angiogenic mitogen, *Science* 246 (1989) 1306–1309.
- [32] J. Folkman, Angiogenesis in cancer, vascular, rheumatoid and other disease, *Nat. Med.* 1 (1995) 27–31.
- [33] J. Folkman, The role of angiogenesis in tumor growth, *Semin. Cancer Biol.* 3 (1992) 65–71.
- [34] W.D. Thompson, K.J. Shiach, R.A. Fraser, L.C. Mcintosh, J.G. Simpson, Tumours acquire their vasculature by vessel incorporation, not vessel ingrowth, *J. Pathol.* 151 (1987) 323–332.
- [35] J. Folkman, Tumor angiogenesis: therapeutic implications, *N. Engl. J. Med.* 285 (1971) 1182.
- [36] J. Yang, D. Yang, Y. Sun, B. Sun, G. Wang, J.C. Trent, et al., Genetic amplification of the vascular endothelial growth factor (VEGF) pathway genes, including VEGFA, in human osteosarcoma, *Cancer* 117 (2011) 4925–4938.
- [37] M.M. Mcmanus, K.R. Weiss, D.P.M. Hughes, Understanding the role of Notch in osteosarcoma, *Adv. Exp. Med. Biol.* 804 (2014) 67–92.
- [38] Q. Wang, Identification of biomarkers for metastatic osteosarcoma based on DNA microarray data, *Neoplasma* 62 (2015) 365.
- [39] M.M. Cooney, W.V. Heeckeren, S. Bhakta, O.S.C. Remick, Drug insight: vascular disrupting agents and angiogenesis—novel approaches for drug delivery, *Nat. Clin. Pract. Oncol.* 3 (2006) 682–692.
- [40] P. Kunz, J. Fellenberg, L. Moskovszky, Z. Sápi, T. Krenacs, I. Machado, et al., Improved survival in osteosarcoma patients with atypical low vascularization, *Ann. Surg. Oncol.* 22 (2015) 489.
- [41] L. Zitvogel, A. Tesniere, G. Kroemer, Cancer despite immunosurveillance: immunoselection and immunosubversion, *Nat. Rev. Immunol.* 6 (2006) 715–727.
- [42] G.P. Dunn, C.M. Koebel, R.D. Schreiber, Interferons, immunity and cancer immunoeediting, *Nat. Rev. Immunol.* 6 (2006) 836–848.
- [43] M.J. Smyth, G.P. Dunn, R.D. Schreiber, Cancer immunosurveillance and immunoeediting: the roles of immunity in suppressing tumor development and shaping tumor immunogenicity, *Adv. Immunol.* 90 (2006) 1.
- [44] V. Shankaran, H. Ikeda, A.T. Bruce, J.M. White, P.E. Swanson, L.J. Old, et al., IFN[gamma] and lymphocytes prevent primary tumour development and shape tumour immunogenicity, *Nature* 410 (2001) 1107.
- [45] K.E.D. Visser, A. Eichten, L.M. Coussens. De Visser, K.E. Eichten, A. Coussens LMParadoxical, roles of the immune system during cancer development, *Nat. Rev. Cancer* 6 (2006) 24–37.
- [46] E.P. Buddingh, M.L. Kuijjer, R.A. Duim, H. Bürger, K. Agelopoulos, O. Myklebost, 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.: Off. J. Am. Assoc. Cancer Res.* 17 (2011) 2110–2119.
- [47] R.A. Mitchell, Mechanisms and effectors of MIF-dependent promotion of tumour-igenesis, *Cell. Signal.* 16 (2004) 13–19.
- [48] W. Es, K.R. Flaherty, S. Carskadon, A. Brant, M.D. Iannettoni, J. Yee, et al., Macrophage migration inhibitory factor and CXC chemokine expression in non-small cell lung cancer: role in angiogenesis and prognosis, *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 9 (2003) 853–860.
- [49] E.H.M.D. Takashi, K.D.M.D. Dipok, N.E.-a.M.D. Osama, H.M.D. Yoshitaka, A. Yamanoi, et al., Overexpression of macrophage migration inhibitory factor induces angiogenesis and deteriorates prognosis after radical resection for hepatocellular carcinoma, *Cancer* 103 (2005) 588–598.
- [50] I. Han, R.L. Mi, K.W. Nam, J.H. Oh, K.C. Moon, H.S. Kim, Expression of macrophage migration inhibitory factor relates to survival in high-grade osteosarcoma, *Clin. Orthop. Relat. Res.* 466 (2008) 2107.
- [51] V.V. Ternovoi, D.T. Curiel, B.F. Smith, G.P. Siegal, Adenovirus-mediated p53 tumor suppressor gene therapy of osteosarcoma, *Lab. Invest.* 86 (2006) 748–766.
- [52] B.D. Ragland, W.C. Bell, R.R. Lopez, G.P. Siegal, Cytogenetics and molecular biology of osteosarcoma, *Lab. Invest.* 82 (2002) 365–373.
- [53] M.D. Vesely, M.H. Kershaw, R.D. Schreiber, M.J. Smyth, Natural Innate and Adaptive Immunity to Cancer, *Annu. Rev. Immunol.*, 29(1):235.
- [54] R.D. Schreiber, L.J. Old, M.J. Smyth, Cancer immunoeediting: integrating immunity's roles in cancer suppression and promotion, *Science* 331 (2011) 1565.