Meta-analysis of global gene-expression profiles identify molecular signatures for histological subtypes of sarcomas

Zhiwei Qiao1, Cuneyd Parlayan1†, Shigeru Saito2* and Tadashi Kondo1**

1 Division of Rare Cancer Research, National Cancer Center Research Institute
2 Data Science Lab, OPT Holding, Inc.
† Present address: Department of Biomedical Engineering School of Engineering and Natural Science & Regenerative Restorative Medicine Research Center (REMER), Istanbul Medipol University, Istanbul, Turkey

(Received September 6, 2018; Accepted October 25, 2018)

SUMMARY

Sarcomas are rare mesenchymal malignancies and comprise over 50 histological subtypes. Sarcomas are not well studied because the number of cases of individual sarcoma is low. The utilization of public data, such as gene expression data, may allow for improvement in the novel discovery of sarcoma. In this study, to gain insight into histological subtypes of sarcoma from a public database, we performed a meta-analysis of the gene-expression profiles by surveying the data deposited in the Gene Expression Omnibus database from 2001 to 2014. The gene-expression data for 10 sarcoma subtypes and the gene-expression profiles for 1002 cases were selected for comparative analysis. Genes with histology-oriented molecular signatures were identified, and the results were verified by functional validation using gene ontology analysis. Pathway analysis suggested the existence of differential biological processes among sarcoma subtypes. Furthermore, as an application of the sarcoma gene expression datasets used in this study, we investigated the gene expression patterns of the targets of pazopanib to predict the response of sarcoma to pazopanib. We found that the gene expression distribution patterns of targets of pazopanib were without distinction among 10 subtypes of sarcoma. Taken together, we identified the tissue-specific genes of 10 subtypes of sarcoma by bioinformatics analysis; our results demonstrated the utility of sarcoma datasets in public databases and provide valuable information for future rare cancer research.

Key words: sarcoma, meta-analysis, gene expression profiles, molecular signatures

INTRODUCTION

Sarcomas display a wide range of histological appearances, for which over 50 distinct subtypes have been defined1). Sarcomas represent a rare and heterogeneous group of malignant tumors, accounting for less than 1% of all malignancies. Although sarcoma subtypes can be distinguished in terms of histopathology, molecular signature, histological grade, and primary site, the boundaries between several diagnostic groups are still vague2-4). The histological diagnosis of sarcoma has been hampered by the low incidence of these tumors5-7). Indeed, the discrepancy between the original histologic assessment and that of a subsequent expert review reached up to 25% of the cases examined8,9). Because the histological subtypes exhibit different clinical behaviors and responses to treatments10), histology-specific treatments can be developed11). Thus, in addition to histological observations, the identification of the genetic traits and molecular pathways underlying histological appearances will provide valuable insight into novel therapeutic modalities for sarcomas.

Molecular biomarkers have been developed to support the histological diagnosis of sarcomas. Lai X et al. identified biomarkers to distinguish low-grade chondrosarcoma from enchondroma12). Valente et al. reported the utility of TLE1 in diagnosing synovial sarcoma13). Ito et al. showed that reduced SMARCB1 expression had diagnostic utility in synovial sarcoma14). These findings will benefit patients with sarcoma after extensive validation studies have established their clinical utility. (Previous data have also demonstrated
distinctive gene-expression patterns among the different histological subtypes of sarcomas\textsuperscript{15–17}, indicating the possibility of identifying subtype-specific genes that can be developed as biomarkers. However, as sarcomas are rare and the number of clinical samples available for research purposes is generally limited, it is difficult to generate conclusive results at a single institution.

The limited number of clinical samples also makes it problematic to share samples among multiple institutions. Alternatively, combining the publicly available gene-expression datasets obtained in multiple studies is a worthwhile challenge\textsuperscript{18–21}. For example, Yang, \textit{Z. et al.} examined 8 Gene Expression Omnibus (GEO) datasets and identified 979 genes with differential expression between osteosarcoma and normal tissues. The differentially expressed genes had distinct biological functions, which explained the unique characteristics of osteosarcoma\textsuperscript{22}. Meta-analysis of gene-expression profiles was also performed using 4 microarray datasets and 2 serial analyses of gene expression (SAGE) datasets\textsuperscript{23}. This analysis resulted in the identification of regulatory pathways of tumor growth in rhabdomyosarcoma\textsuperscript{23}. Hancock et al. reported gene sets associated with EWS-FLI dysregulation in Ewing’s sarcoma model systems\textsuperscript{24}. Villacis et al. investigated the publicly available datasets and found that SRC has diagnostic utility in differentiating leiomyosarcoma from undifferentiated pleomorphic sarcoma\textsuperscript{25}. All these studies suggested the possible utility of meta-analysis of sarcomas using publicly available datasets. However, because of the limited number of samples and sarcoma subtypes identified in previous studies, meta-analysis of sarcomas was not fully pursued.

With this premise in mind, we sought to examine all sarcoma-related gene data deposited in GEO from 2001 to 2014, enabling us to identify gene expression profiles unique to sarcoma subtypes. We examined molecular pathways enhanced in individual sarcoma subtypes. We also used the gene expression datasets to predict the response of sarcoma to pazopanib treatment.

**MATERIALS AND METHODS**

1. Search strategy and selection criteria

We conducted a gene-expression dataset search using the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). Based on the World Health Organization’s classification of tumors of soft tissue and bone\textsuperscript{26}, we searched for gene expression data obtained from patients with the 32 sarcoma subtypes from January 2001 to December 2014. All of the keywords used for searching were listed in Supplementary Table 1. Information associated with the gene-expression data in the GEO database, as well as related publications were used to confirm the characteristics of the samples. Studies were eligible for the meta-analysis if they met the following inclusion criteria: 1) experiment type was expression profiling by array, 2) samples were clinical specimens, 3) data were generated using the Affymetrix Human Genome U133A Array (GPL96) or Affymetrix Human Genome U133 Plus 2.0 Array platform (GPL570) 4) datasets contained original data.

2. Data extraction

The two investigators independently extracted relevant data and reached a consensus on all items. The following information was curated for each data set: GEO accession number, tumor type, publication, DNA microarray platform, number of cases, references, and gene-expression data.

3. Bioinformatic analysis

Gene-expression data analysis was performed using R software and software packages from the Bioconductor Project as follows\textsuperscript{28,29}. The DNA-microarray data were normalized using MAS5.0 with the Bioconductor affy package\textsuperscript{29} and then by the global median centering method. Common probe sets (22,227 probes) among GPL96 and GPL570 were selected and used for further analysis. To eliminate batch effects between different DNA-microarray platforms, we employed the COMBAT algorithm\textsuperscript{29}, which was based on the empirical Bayes method, using the Bioconductor in Silico Merging package\textsuperscript{30}. For unsupervised classification, a z-score transformation was conducted with all DNA-microarray data. The normalized data were subjected to hierarchical clustering by determining the Euclidean distance and by Ward’s method. To obtain representative gene-expression data for each tissue subtype, the median gene-expression levels in each tissue type were calculated for all probes used. To identify tissue-specific gene expression, the ROKU method\textsuperscript{31}, which is based on determining the Shannon entropy to measure the specificity of gene expression\textsuperscript{32}, was employed and performed using Bioconductor ROKU package with default parameters. Genes that were present or absent only in a single tissue were defined as tissue-specific probes, according to the Shannon entropy calculation.

Gene expression data of nine pazopanib target genes (VEGFR1, VEGFR2, VEGFR3, FGFR1, FGFR2, FGFR3, PDGFR, PDGFRB, and KIT) were extracted from expression data of 1002 samples used in this study. Hierarchical clustering was performed using Pearson’s correlation distance and Ward’s method to verify the distribution of targets genes of pazopanib among 10 subtypes of sarcoma.

4. Pathway and gene ontology enrichment analysis

For each gene-expression group for each tumor subtype, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotations were inferred using Database for Annotation, Visualization and Integration Discovery (DAVID) software (http://david.abcc.ncifcrf.gov/). The KEGG database was used to classify gene sets correlated with tumor subtypes into respective pathways\textsuperscript{30}. The significance of
gene enrichment was indicated by the p-value for each category, and process groups were considered significant with p < 0.05. KEGG analysis results were plotted using the R package “tree map”\textsuperscript{34}. To further understand the functions of the subset of genes identified, we performed Gene Ontology (GO) enrichment analysis, using the hyper geometric test to detect enriched functional attributes based on DAVID software (p < 0.05).

**RESULTS**

1. Process of gene-expression data selection; inclusion and exclusion criteria

Database searches yielded a total of 714 studies and 25 studies met the inclusion criteria. The process used to select the studies included the probe datasets for further analysis was illustrated in Fig. 1. We searched for the probe data of 32 sarcoma subtypes (Supplementary Table 1). In the GEO database, we found DNA-microarray data for 21 out of the 32 sarcoma subtypes, which were derived from 2253 samples (Supplementary Table 2). All GSE files and published information used in this study are summarized in Supplementary Table 3. Over half of the deposited data were generated using the Affymetrix GPL96 and GPL570 platforms, which were based on the use of Affymetrix Human Genome U133 series. Thus, we decided to use only the probe data obtained on either of these 2 platforms for further analysis. Probe datasets without original data (.CEL files) were excluded from this study. Consequently, we included 10 tissue types and 1002 raw data files (.CEL files) for further meta-analysis. The selected samples included 25 alveolar soft part sarcoma samples, 34 clear cell sarcoma samples, 183 Ewing’s sarcoma samples, 36 fibrosarcoma samples, 83 gastrointestinal stromal tumor (GIST) samples, 117 leiomyosarcoma samples, 356 liposarcoma samples (including 161 dedifferentiated, 69 myxoid, 54 pleomorphic, and 72 well-differentiated liposarcoma), 75 osteosarcoma samples, 59 malignant rhabdoid tumor samples, and 34 synovial sarcoma samples (Table 1).

2. Identification of subtype-specific gene expression profiles

We analyzed the overall features of the probe data from 14 tissue types (including 4 different subtypes of liposarcoma) and 1002 samples. Unsupervised hierarchical clustering was used to explore overall relationships occurring among samples and the underlying gene-expression features (Fig. 2A). The results of the unsupervised sample classification based on the probe data did not correlate with the histological classification of sarcomas (Fig. 2A). These observations led us to identify probe sets unique to sarcoma subtypes. We per-

---

Fig. 1. Flow diagram of the process used to select microarray datasets for the meta-analysis.

Publically available raw microarray data from sarcoma samples deposited in the GEO database were selected for meta-analysis.
formed a 1-versus-all comparison using the ROKU method and identified probes with differential intensities unique to specific histologies (Supplementary Table 4). Redundant probes were eliminated and the remaining selected probes were subjected to supervised clustering analysis (Fig. 2B). We found that the wide distribution of selected probe-intensity patterns was consistent with the prevalence of the tumor type.

The number of probes showing signal intensities unique to a particular histology was diverse among the sarcomas. We hypothesized that the heterogeneity of molecular backgrounds may correlate negatively with the number of tissue-specific genes, due to 2 mechanisms. Firstly, the small number of unique probes may indicate the complexity of the molecular background in sarcoma subtypes. For example, although liposarcomas consist of multiple subtypes, which may have different biological characteristics, these data were merged and subjected to hierarchical clustering in this study. Then, comparative analysis resulted in identification of the minimal number of unique genes in liposarcoma. Second, in addition to the possible heterogeneity of molecular backgrounds, in general, the more samples included in the gene-expression study, the less group-specific genes that were identified, because the degree of distribution follows a power law. To explore this possibility, we plotted the logarithm of the number of probes as a function of the logarithm of the number of samples (Fig. 3). We found a negative correlation between the number of samples and the number of specific genes from alveolar soft part sarcoma (ASPS) to liposarcomas. The plotted data also suggested that GIST and Ewing’s sarcoma samples may have relatively homogeneous molecular backgrounds because they were plotted above the average curve, whereas fibrosarcoma may have higher heterogeneity because it was plotted below the average curve (Fig. 3).

Table 1. Summary of gene expression data and platform

| Tissue type                                      | Affymetrix Human Genome U133 Plus 2.0 Array (GPL570) | Affymetrix Human Genome U133A Array (GPL96) | Total |
|-------------------------------------------------|-----------------------------------------------------|---------------------------------------------|-------|
| Alveolar Soft Part Sarcoma (ASPS)               | 25                                                  | 0                                           | 25    |
| Clear cell sarcoma                              | 1                                                   | 33                                          | 34    |
| Ewing sarcoma                                   | 183                                                 | 0                                           | 183   |
| Fibrosarcoma                                    | 0                                                   | 36                                          | 36    |
| Gastrointestinal stromal tumor (GIST)           | 83                                                  | 0                                           | 83    |
| Leiomyosarcoma                                  | 91                                                  | 26                                          | 117   |
| Liposarcoma                                     | 91                                                  | 265                                         | 356   |
| Osteosarcoma                                    | 27                                                  | 48                                          | 75    |
| Malignant rhabdoid tumor                        | 39                                                  | 20                                          | 59    |
| Synovial sarcoma                                | 34                                                  | 0                                           | 34    |
| Total                                           | 574                                                 | 428                                         | 1002  |

3. Pathway analysis of the identified genes in the subtypes of sarcoma

Tree maps of KEGG pathways were created for each sarcoma subtype (Fig. 4). The gene annotations used in the tree-maps are summarized in Supplementary Table 5. Each treemap included all KEGG pathways with adjusted p-value < 0.05, colored according to their enrichment score and sized according to the number of genes in that pathway. Among the 10 subtypes included in this study, unique KEGG pathways were identified in 8 subtypes with statistical significance (p < 0.05). The enriched KEGG pathways were quite different among the sarcomas. GIST was characterized by genes associated with the phosphotyidylinositol-signaling pathway, including PIK3CG, INPP1, PLCE1, PLCB4, PLCG1, ITPKB, PI4KB, OCRL, PIP4K2A, INPP5B, INPP5A, and PIK3R1 (Fig. 4A). The neuroactive ligand-receptor interaction pathway, ribosome pathway, ubiquitin mediated proteolysis pathway, focal adhesion pathway, extracellular matrix (ECM) receptor interaction pathway, neuroactive ligand receptor pathway, and pathway in systemic lupus erythematosus, were dominant in leiomyosarcoma, Ewing’s sarcoma, clear cell sarcoma, ASPS, osteosarcoma, rhabdoid sarcoma, and synovial sarcoma, respectively (Fig. 4, B–H).

DISCUSSION

In this study, we performed a meta-analysis of the gene expression data from sarcomas. To our knowledge, this is the first report describing the investigation of all major sarcoma subtypes and identifying sarcoma subtype-specific genes through a meta-analysis. Because there are no other gene expression datasets available for sarcoma for validation, we validated the subtype-specific genes identified by functional analysis. As shown in the results, we performed the gene ontology analysis, and analyzed biological processes of three types of sarcoma. In GIST, we found that three biological processes were significantly associated with the neural system. In ASPS, two
In some cases, with previously published data\textsuperscript{16}. For example, enhanced signaling through the Wnt pathway was identified in synovial sarcoma\textsuperscript{39–41}, as was enhancement of the phosphatidylinositol-signaling pathway in GIST\textsuperscript{42}, the focal adhesion pathway in leiomyosarcoma\textsuperscript{43}, and the ECM-receptor interaction pathway in osteosarcoma\textsuperscript{22}. Although biological processes were significantly associated with the blood vessels. In osteosarcoma, two biological processes were associated with the bone tissue. All of these results supported the reliability of the tissue-specific genes identified in this study.

The pathway results of our study are also concordant, in some cases, with previously published data\textsuperscript{46}. For example, enhanced signaling through the Wnt pathway was identified in synovial sarcoma\textsuperscript{39–41}, as was enhancement of the phosphatidylinositol-signaling pathway in GIST\textsuperscript{42}, the focal adhesion pathway in leiomyosarcoma\textsuperscript{43}, and the ECM-receptor interaction pathway in osteosarcoma\textsuperscript{22}. Although

**Fig. 2.** Cluster analysis of 10 sarcoma samples.

Unsupervised hierarchical clustering of 10 sarcoma subtypes was performed using 1002 samples and 22,277 genes (A). The genes unique to each sarcoma subtype were combined and subjected to hierarchical analysis (B). Each column represents a sample, and each line represents gene-expression data. Genes with relatively high expression levels are represented in red, while genes with relatively low expression levels are represented in green.
CONCLUSIONS

We performed a meta-analysis of gene-expression data across 10 sarcoma subtypes and detected subtype-specific gene expression. Our results demonstrated the utility of examining sarcoma datasets from public databases. Our results also suggested the possible value of publicly available datasets in identifying genes for clinical and pathological parameters in sarcomas.

ACKNOWLEDGEMENTS

This research was supported by the Practical Research for Innovative Cancer Control Program 303 (15ck0106089h0002) from the Japan Agency for Medical Research and Development (AMED).

ABBREVIATIONS

DAVID, Database for Annotation, Visualization and Integration Discovery; GEO, Gene Expression Omnibus; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GIST, gastrointestinal stromal tumor

SUPPORTING INFORMATION

Supplementary Table 1. List of sarcomas examined in this study.
Supplementary Table 2. Summary of sarcoma subtypes and samples in this study.
Supplementary Table 3. Summary of supplemental information.

we also found novel pathways for each sarcoma subtype, these results should be further validated using additional data from independent samples.

As an application of our datasets, we examined the targets of gene expression datasets to predict the response to pazopanib. We found that gene expression patterns were similar among 1002 samples, suggesting that it is difficult to predict the response to pazopanib using gene expression data. The protein expression level or kinase activity level of targets might need to be used to predict the response to pazopanib treatment.

Although meta-analysis of sarcoma data appears to be promising, several inherent limitations should be considered. Firstly, the results of the meta-analysis are dependent on the reliability of the original data\(^{44-46}\), which was not validated in this study. The use of a large number of datasets may partially overcome this issue, and we began our study using all relevant GEO microarray data deposited between 2001 and 2014. Secondly, only genes included in the platforms used to generate the data could be used for analysis; thus, the results may have been biased due to the popularity of the platform. In this study, we focused on DNA-microarray data obtained using the Affymetrix Human Genome U133 series. The data content associated with a given DNA-microarray format should be considered for a better interpretation of the meta-analysis data.

The results of gene-enrichment analysis of several types of sarcomas are shown for GIST (a), leiomyosarcoma (b), Ewing’s sarcoma (c), clear cell sarcoma (d), ASPS (e), osteosarcoma (f), rhabdoid tumor (g), and synovial sarcoma (h). Gene ontology-category enrichments of differentially regulated genes, based on the KEGG pathway. The box sizes represent the numbers of genes in each category, and the color represents the enrichment score. All process groups were considered significant when \( p < 0.05 \).
REFERENCES

1) Murphey MD. World Health Organization classification of bone and soft tissue tumors: modifications and implications for radiologists. Semin Musculoskelet Radiol. 2007;11:201–214.

2) Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin. 2014;64:9–29.

Supplementary Table 4. Specific probes and gene symbols identified in 10 subtypes of sarcoma.
Supplementary Table 5. Genes included in the KEGG pathways of 10 subtypes of sarcoma.

The logarithm of the number of probes is plotted as a function of the logarithm of the number of samples.
3) Norberg SM, Movva S. Role of genetic and molecular profiling in sarcomas. Current Treatment Options in Oncology. 2015;16:24.
4) Doyle LA. Sarcoma classification: an update based on the 2013 World Health Organization Classification of Tumors of Soft Tissue and Bone. Cancer. 2014;120:1763–1774.
5) Herzog CE. Overview of sarcomas in the adolescent and young adult population. J Pediatr Hematol Oncol. 2005;27:215–218.
6) Demicco EG. Sarcoma diagnosis in the age of molecular pathology. Adv Anat Pathol. 2013;20:264–274.
7) Bode-Lesniewska B. [Cytologic diagnosis of sarcoma]. Pathologie. 2011;32:14–23.
8) Presant CA, Russell WO, Alexander RW, Fu YS. Soft-tissue and bone sarcoma histopathology peer review: the frequency of disagreement in diagnosis and the need for second pathology opinions. The Southeastern Cancer Study Group experience. J Clin Oncol. 1986;4:1658–1661.
9) Shiraki M, Enterline HT, Brooks JJ, et al. Pathologic analysis of advanced adult soft tissue sarcomas, bone sarcomas, and mesotheliomas. The Eastern Cooperative Oncology Group (ECOG) experience. Cancer. 1989;64:484–490.
10) Blay JY, Sleijfer S, Schoffski P, et al. International expert opinion on patient-tailored management of soft tissue sarcomas. Eur J Cancer. 2014;50:679–689.
11) Sethi TK, Keedy VL. Histology-specific uses of tyrosine kinase inhibitors in non-gastrointestinal stromal tumor sarcomas. Current Treatment Options in Oncology. 2016;17:11.
12) Lai X, Chen S. Identification of novel biomarker candidates for immunohistochemical diagnosis to distinguish low-grade chondrosarcoma from enchondroma. Proteomics. 2015;15:2358–2368.
13) Valente AL, Tull J, Zhang S. Specificity of TLE1 expression in unclassified high-grade sarcomas for the diagnosis of synovial sarcoma. Appl Immunohistochem Mol Morphol. 2013;21:408–413.
14) Ito J, Asano N, Kawai A, Yoshida A. The diagnostic utility of reduced immunohistochemical expression of SMARCB1 in synovial sarcomas: a validation study. Human Pathology. 2016;47:32–37.
15) Baird K, Davis S, Antonescu CR, et al. Gene expression profiling of human sarcomas: insights into sarcoma biology. Cancer Res. 2005;65:9226–9235.
16) Segal NH, Pavlidis P, Antonescu CR, et al. Classification and subtype prediction of adult soft tissue sarcoma by functional genomics. Am J Pathol. 2003;163:691–700.
17) West RB, van de Rijn M. The role of microarray technologies in the study of soft tissue tumours. Histopathology. 2006;48:22–31.
18) Takahashi A, Nakayama R, Ishibashi N, et al. Analysis of gene expression profiles of soft tissue sarcoma using a combination of knowledge-based filtering with integration of multiple statistics. PLoS One. 2014;9:e106801.
19) Wang Z, He ML, Zhao JM, Qing HH, Wu Y. Meta-analysis of associations of the ezrin gene with human osteosarcoma response to chemotherapy and prognosis. Asian Pac J Cancer Prev. 2013;14:2753–2758.
20) Nakayama R, Nemoto T, Takahashi H, et al. Gene expression analysis of soft tissue sarcomas: characterization and reclassification of malignant fibrous histiocytoma. Mod Pathol. 2007;20:749–759.
21) Mills AM, Beck AH, Montgomery KD, et al. Expression of subtype-specific group 1 leiomyosarcoma markers in a wide variety of sarcomas by gene expression analysis and immunohistochemistry. Am J Surg Pathol. 2011;35:583–589.
22) Yang Z, Chen Y, Fu Y, et al. Meta-analysis of differentially expressed genes in osteosarcoma based on gene expression data. BMC Med Genet. 2014;15:80.
23) Romualdi C, De Pita C, Tombolan L, et al. Defining the gene expression signature of rhabdomyosarcoma by meta-analysis. BMC Genomics. 2006;7:287.
24) Hancock JD, Lessnick SL. A transcriptional profiling meta-analysis reveals a core EWS-FLI1 gene expression signature. Cell Cycle. 2008;7:250–256.
25) Villacis RA, Silveira SM, Barros-Filho MC, et al. Gene expression profiling in leiomyosarcomas and undifferentiated pleomorphic sarcomas: SRC as a new diagnostic marker. PLoS One. 2014;9:e102281.
26) Ihaka R RG. A language for data analysis and graphics. J Comput Graph Stat. 1996;5:299–314.
27) Reimers M, Carey VJ. Bioconductor: an open source framework for bioinformatics and computational biology. Methods in Enzymology. 2006;411:119–134.
28) Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20:307–315.
29) Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8:118–127.
30) Taminau J, Meganck S, Lazar C, et al. Unlocking the potential of publicly available microarray data using inSilicoDb and inSilicoMerging R/Bioconductor packages. BMC Bioinformatics. 2012;13:335.
31) Kadota K, Ye J, Nakai Y, Terada T, Shimizu K, ROKU: a novel method for identification of tissue-specific genes. BMC Bioinformatics. 2006;7:294.
32) Fuhrman S, Cunningham MJ, Wen X, Zweiger G, Seilhamer JJ, Somogyi R. The application of shannon entropy in the identification of putative drug targets. Bio Systems. 2000;55:5–14.
33) Altermann E, Klaenhammer TR. PathwayVoyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. BMC Genomics. 2005;6:55.
34) Tyanova S, Albrechtsen R, Kronqvist P, Cox J, Mann M, Geiger T. Proteomic maps of breast cancer subtypes. Nat Commun. 2016;7:10259.
35) Small CG. The statistical theory of shape. New York: Springer-Verlag; 1996. 227 p.
36) Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL. The origins of power-law distribution of the degree of scale-free networks. Proc Nat Acad Sci USA. 2000;98:11159–11163.
degree distribution in the heterogeneity of human activity in social networks. Scientific Reports. 2013;3:1783.

39) Trautmann M, Sievers E, Aretz S, et al. SS18-SSX fusion protein-induced Wnt/beta-catenin signaling is a therapeutic target in synovial sarcoma. Oncogene. 2014;33:5006–5016.

40) Barham W, Frump AL, Sherrill TP, et al. Targeting the Wnt pathway in synovial sarcoma models. Cancer Discov. 2013;3:1286–1301.

41) Cironi L, Petricevic T, Fernandes Vieira V, et al. The fusion protein SS18-SSX1 employs core Wnt pathway transcription factors to induce a partial Wnt signature in synovial sarcoma. Sci Rep. 2016;6:22113.

42) Daniels M, Lurkin I, Pauli R, et al. Spectrum of KIT/PDGFRα/BRAF mutations and Phosphatidylinositol-3-Kinase pathway gene alterations in gastrointestinal stromal tumors (GIST). Cancer Lett. 2011;312:43–54.

43) Owens LV, Xu L, Craven RJ, et al. Overexpression of the focal adhesion kinase (p125FAK) in invasive human tumors. Cancer Res. 1995;55:2752–2755.

44) Hong F, Breitling R. A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. Bioinformatics. 2008;24:374–382.

45) Guilhamon P, Eskandarpour M, Halai D, et al. Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2. Nat Commun. 2013;4:2166.

46) Yu XW, Wu TY, Yi X, et al. Prognostic significance of VEGF expression in osteosarcoma: a meta-analysis. Tumour Biol. 2014;35:155-60.