Bioinformatics Analysis of the Biological Properties of Ewing Sarcoma

Luchang Chen¹, Huifang Zeng¹, Wujia Yang¹, Haidong Zhou¹, Changtai Luo¹, Dong Luo¹, Zhenjing Si¹, Wei Wang¹, Jihua Wei²*

¹Graduate School of Youjiang Medical University for Nationalities, Baise, China
²Department of Sports Medicine, Baidong Hospital Affiliated to Youjiang Medical College for Nationalities, Baise, China

Email: *1261290953@qq.com


Received: May 21, 2024
Accepted: June 21, 2024
Published: June 24, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

Purpose: Bioinformatics-based approach to screen and analyze differentially expressed genes associated with the biological characteristics of Ewing sarcoma. Means: The GSE17674 dataset was selected for analysis, obtained by data retrieval based on the GEO public database. The R language limma tool-kit was used to screen DEmRNAs. After the data were normalized, the Metascape online analysis software and the R language clusterProfiler package were used to analyze the GO function and KEGG pathway enrichment of DEmRNAs lines, respectively. The string database was selected for PPI analysis, and the results were imported into Cytoscape software to derive the core modules and predicted core genes. The genes selected above were analyzed for tissue localization specificity. Results: Through the analysis of GSE17674, differentially expressed genes were screened out, and GO and KEGG analyses were performed on the differentially expressed genes. The GO functional enrichment analysis was mainly enriched in the process of muscle system, muscle contraction, myocyte development, contractile fibers, myogenic fibers, myofibers, myofibrillar segments, actin binding, structural composition of muscle, and actin filament binding. KEGG pathway analysis showed that the core pathways associated with the development of ES were the core genes for myocardial contraction, congestive cardiomyopathy, and hypertrophic cardiomyopathy. Five Hub genes were obtained based on Cytoscape prediction. Tissue localization specificity analysis of Hub genes was performed, and a total of 2 Hub genes with tissue specificity were screened; MYH6 was specifically expressed in cardiac cells and MYL1 was specifically expressed in skeletal muscle cells. Conclusions: The differential genes screened will help to understand the molecular mechanisms underlying the highly invasive and metastasis-prone biological characteristics of ES, as well as provide new ideas for clinical drug-targeted treatment of ES.
Keywords

Ewings Sarcoma, Myosin, Bioinformatics Analysis, Targeted Genes

1. Backgrounds

Ewings sarcoma (ES) is a malignant aggressive bone or soft tissue tumor that occurs in children, adolescents, and young adults, with a peak incidence at age 15 [1]. It accounts for approximately 10% - 15% of all primary bone tumors in children, adolescents, and young adults and is the second most common primary bone tumor after osteosarcoma [2]. Studies have shown that in Europe [1], 7.5 cases per million children between the ages of 10 and 19 suffer from the tumor each year. And, the prevalence of ES is slightly higher in men than in women, with a sex ratio of 3:2. ES most commonly affects the bones, especially the humerus, ribs, pelvis and femur [3]. However also 20% - 30% of ES occurs in the patient’s soft tissues, such as the chest wall or pleural cavity. ES is characterized by rapid tumor growth and active metastasis, and is highly susceptible to metastasis. The most common sites of metastasis are the lungs, bones, and also the spinal cord. Ewing sarcoma is characterized by chromosomal translocations, the most common being the t(11;22)(q24;q12) chromosome abnormality, which occurs in 85% of cases. This leads to the fusion of the FET family gene EWSR1 with the ETS family gene FLI1 [4], producing the EWS-FLI1 fusion protein. The EWS-FLI1 fusion protein, an aberrant transcription factor, is critical for the development, maintenance and progression of Ewing sarcoma [5]. It plays a key role in controlling the expression of many target genes and coordinating the oncogenic process of malignant transformation of precursor cells into cancerous cells.

Previous treatments for Ewing sarcoma, including aggressive neoadjuvant chemotherapy and adjuvant chemotherapy combined with surgery and/or radiation, have improved long-term survival in patients with limited disease, with a 5-year survival rate of more than 70% and a 10-year survival rate of approximately 30%. However, once ES tumor cells metastasize or recur, the 5-year survival rate drops dramatically to 25% [2] [6]. This has not improved over the decades, so there is an urgent need to explore more effective therapeutic targets to reduce the incidence of metastatic and recurrent events in ES patients.

Therefore, in this study, we performed data search based on the GEO public database and chose the GSE17674 dataset as the object of analysis, and gained insight into the key differentially expressed genes by screening differential genes → GO and KEGG enrichment analysis → construction of protein-protein interaction (PPI) network → screening of Hub genes for their biological functions, with a view to screening key molecules for the biological characterization of Ewing sarcoma.
2. Materials and Methods

2.1. General Materials

This study was based on the GEO public database for data retrieval, which was first proposed in 2012 by Benisch et al. [7]. The GSE17674 dataset was selected for analysis, and the set platform number of the data patients was GPL570. This data contains 62 samples, of which 44 were Ewing sarcoma tumor tissue from patients and 18 were normal muscle tissue.

2.2. Expression Analysis of Differential Genes

In this study, we performed differential expression analysis of Ewing sarcoma mRNA microarray data using the R language limma package, and set the test statistic $P < 0.05$ and the log absolute value of fold change (FC) $|\log FC| > 1$ as the screening conditions to screen the differentially expressed mRNAs (DEmRNAs) of Ewing sarcoma and normal muscle tissue.

2.3. GO Function and KEGG Signaling Pathway Enrichment Analysis

In this study, DEmRNAs were analyzed for GO functional enrichment and KEGG signaling pathway enrichment based on the R language clusterProfiler package.

2.4. PPI and Core Module Analysis

Differential genes were analyzed by STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) online tool to obtain PPI maps, the images were imported into Cytoscape 3.9.1 software, and the plugins Cytohubba and MCODE were used to search for the Ewing sarcoma-related Hub genes and Hub modules.

2.5. Hub gene localization

Tissue localization of the above genes was performed by the online database BioGPS (http://biogps.org/#goto=welcome), and the highest expression between tissues was more than twice the second expression was meaningful, indicating that this gene can be used as an ES expression marker gene.

3. Results

3.1. Acquisition of DEmRNAs

In this study, GSE17674 was analyzed using the R language limma package, and the gene expression of the database samples remained largely consistent, indicating that the data is suitable for the next step in the analysis, Figure 1.

Screening for differential genes using the R language limma package, Figure 2, Figure 3 and Figure 4.

3.2. GO and KEGG Enrichment Analysis

To understand the functions of up- and down-regulated DEmRNAs, GO and KEGG enrichment analyses of up-regulated DEmRNAs, and down-regulated
DEmRNAs, respectively, were performed in this study. Biological Processes (BP), Cellular Components (CC), Molecular Functions (MF) and signaling pathways (pathway) in KEGG were analyzed separately. From Figure 5, we can learn that: (i) DEmRNAs are enriched in biological processes (BP), molecular functions (MF), and cellular components (CC): For BP, DEmRNAs are mainly enriched in muscle system process, muscle contraction, and muscle cell development. For CC, DEmRNAs were mainly enriched in contractile fiber, myofibril, and sarcomere. For MF, DEmRNAs were mainly enriched in actin binding, structural constituents of muscle, and muscle contraction. For MF, DEmRNAs were mainly enriched in actin binding, structural constituent of muscle, and actin filament binding. (ii) KEGG analysis showed that DEmRNAs were mainly enriched in Cardiac muscle contraction, Dilated cardiomyopathy, and Hypertrophic cardiomyopathy.

Figure 1. GSE17674 dataset box line diagram.

Figure 2. GSE17674 volcano map.
Figure 3. GSE17674 heat map.

Figure 4. Aerial view of GSE17674.
3.3. Protein Mutual Network Construction and Module Analysis of Common Differentially Expressed Genes

The DEmRNAs were analyzed using STRING, and the minimum interaction score was set to 0.7, which resulted in an interaction graph consisting of 204 nodes and 555 connectors. The calculation results of STRING were imported into Cytoscape 3.9.1 software to obtain the PPI network interaction map (Figure 6(a)); the protein interaction network was analyzed with the MCODE plug-in to obtain the core modules of DEmRNAs (Figure 6(b)); and five Hub genes were screened with Cytoscape, Figure 7.
Figure 6. Differential gene PPI network and core modules. (a) A PPI network diagram, (b) Core module.

Note: Red color means high correlation with the disease, lighter color means lower correlation.

Figure 7. Hub genes predicted by Cytoscape.

3.4. Gene Localization Analysis

The core genes predicted by Cytoscape, were imported into the BioGPS database, and the screening criteria were: (i) tissue-specific expression level > 10-fold of the median. And (ii) the second highest expression level was less than one-third of the highest level. A total of 2 Hub genes with tissue-specific expression were screened, MYH6 was specifically expressed in cardiac cells and MYL1 was specifically expressed in skeletal muscle cells, Figure 8.

3.5. Targeted Gene Screening

Five core genes were obtained based on Cytoscape prediction: MYL1, MYH7, MYH6, MYL2, MYH2, of which MYH6 and MYL1 were tissue-specific genes. In summary, a total of five Hub genes closely related to ES pathogenesis were screened, among which MYL1 and MYH7 had the highest relevance.
4. Discussion

Ewing sarcoma is characterized as highly aggressive and highly susceptible to metastasis. In-depth study of the biological characteristics of ES and exploration of the role of differentially expressed genes in the development of ES can provide an important basis for the clinical targeting of ES patients and can reduce the occurrence of recurrence and metastatic events in ES patients after treatment. It is well known that tumor microenvironment (TME) plays an important role in tumor development, progression and metastasis [8]. Mechanical forces within the TME play a key role in driving physiological and pathological processes in tumors and orchestrate their behavior in a variety of biological processes, including cell division, survival, differentiation, and migration [9] [10]. In solid tumors, as the number of tumor and non-tumor cells increases, the pressure within the tumor elevates, leading to higher mechanical forces. These forces transmit signals to the tumor cells, ultimately driving tumor progression [11].

The cytoskeleton plays a key role in the signaling of mechanical forces. Cells are connected to the external environment through their cytoskeleton, which receives external signals that direct cells to undergo invasion and migration [12].

Myosin plays an important role in regulating cytoskeletal organization and dynamics [13]. In humans, the myosin family is organized into 12 distinct...
classes [14]. Of these, class II was the first to be discovered and is the most well characterized class of myosin [15]. Myosin is involved in intracellular functions such as cell migration, adhesion, intracellular transport, signal transduction and tumor suppression [16]. And myosin is also associated with tumor progression and metastasis [17]. Derycke et al. [18] reported that the involvement of myosin IIA in the promotion of EMT, cell invasion and metastasis could be a target for breast cancer therapy. Myosin II may serve as a novel therapeutic target for targeting strategies to inhibit invasion and metastasis of Ewing sarcoma tumor cells. Myosin is a cytoskeletal protein, a hexamer consisting of 2 heavy chains, 2 essential light chains, and 2 regulatory light chains, and is an important component of muscle tissue. Myosin light chain (MLC) is regulated by the MYL gene family and can be classified into two types, a basic light chain (MYL1) and a regulatory light chain (MYL2); myosin heavy chain is regulated by the MYH gene family [19].

Aberrant expression of myosin light chain1 (MYL1) is involved in a variety of pathophysiological processes. Ravenscroft [20] found that MYL1 abnormalities may lead to structural and functional abnormalities in myosin, which in turn may lead to congenital myopathies, suggesting that MYL1 may be a potential target for the diagnosis and treatment of congenital myopathies. Using bioinformatics analysis, Lin Z and Lin Y [21] found that MYL1 may be involved in the development and progression of steroid-induced osteonecrosis of the femoral head by regulating muscle function, suggesting that MYL1 may be a potential therapeutic target for patients with osteonecrosis of the femoral head. MYL1 is also recognized as an important genetic marker for rhabdomyosarcoma and may be involved in the pathogenesis of rhabdomyosarcoma. Studies have shown that MYL1 is highly expressed in rhabdomyosarcoma patients and positively correlates with patient prognosis [22], and that MYL1 is involved in the development and progression of rhabdomyosarcoma by regulating normal muscle differentiation and structural components of muscle. MYL1 may be a molecular target for early diagnosis and treatment of rhabdomyosarcoma. Moreover, MYL1 was shown to be a marker of poor prognosis in head and neck squamous cell carcinoma (HSCC). MYL1 can promote HSCC metastasis associated with tumor immune infiltration in HNSCC, and MYL1 enhances the protein expression levels of EGF and EGFR in HSCC, these effects of MYL1 may be related to the EGF/EGFR signaling pathway [23]. Therefore, MYL1, as a core gene of ES, plays an important role in the development of ES and may be the main gene responsible for the high aggressiveness and susceptibility to metastasis of ES. The MYL1 gene may be a targeted therapy to reduce the recurrence rate and improve the survival rate of ES patients in the future.

The myosin regulatory light chain (MRLC) encoded by the myosin light chain 2 (MYL2) gene is a target of myosin light chain kinase. Myosin regulatory light chains play an important role in the regulation of cellular motility [24]. A key regulatory mechanism of cellular motility is the control of myosin activity, which in
non-muscle cells is determined by phosphorylation of myosin regulatory light chains [25]. Actin contractility is regulated by MRLC diphosphorylation, which generates cellular contractile forces through actin contraction, and these contractile forces are critical for the invasive capacity of cancer cells [26]. It has been shown that phosphorylation of myosin regulatory light chains enhances cancer cell invasion [27], whereas, conversely, inhibition of myosin regulatory light chain phosphorylation suppresses tumor cell growth and invasive capacity [28] [29]. Based on bioinformatics analysis, MYL2 gene was found to be one of the core genes of ES. The high expression of MLL2 in ES tumor cells and the increased phosphorylation of its MRLC may also contribute to its high aggressiveness.

The MYH7 gene is located on chromosome 14 and consists of 41 exons. MYH7 mutations are a common cause of hypertrophic [30] or dilated cardiomyopathy [31], Laing distal myopathy [32], and myosin storage myopathy. It was shown that MYH7 was involved in the immune infiltration of HNSCC [23]. This has led to the possibility that MYH7 may be associated with the prognosis of patients with HNSCC and is also one of the markers of poor prognosis and correlates with tumor staging or grading in HNSCC. MYH7 is one of the cores of ES, and the immune infiltration it participates in may be closely related to the development of ES.

MYH6 expresses one of the two myosin heavy chain isoforms expressed by cardiac muscle, alpha-cardiac myosin heavy chain, which is an important protein component in the formation of cardiac structure and is actively involved in muscle force generation and contraction [33]. It has a potential impact on a wide range of cardiac diseases such as hypertrophic cardiomyopathy, dilated cardiomyopathy, sick sinus node syndrome, atrial septal defects and ventricular septal defects. Recent studies have shown that genetic polymorphisms in MYH6 are associated with the risk of ventricular septal defects [34] [35]. The major proteins interacting with the MYH6 protein are essentially involved in myocardial development. Altered MYH6 protein expression may affect interactions with other proteins, which may have important implications for myocardial development [36]. Thus, low expression of MYH6 protein may lead to defective myocardial formation. Deletion of the MYH6 gene has been reported to be responsible for atrial septal defects [37]. The MYH6 gene is specifically expressed in cardiac cells. However, there are few reports on the role of MYH6 in the development of ES. Further studies are warranted.

MYH2 is one of the isoforms of myosin heavy chain. Mutations in the MYH2 gene are the main cause of MYH2-associated myopathies presenting as ptosis and extraocular muscle paralysis [38]. MYH2 has been shown to be a marker for differentiating squamous cell carcinoma of the head and neck from squamous cell carcinoma of the lung [39]. A new study found [40] that in patients with squamous carcinoma of the head and neck, MYH2 expression was significantly and positively correlated with the level of CD4+ T cell and macrophage infiltration. It suggests that genetic mutations and differential expression of the MYH2 gene may be an
important driver of immune infiltration by CD4+ T cells and macrophages. This suggests that genetic mutations and differential expression of the MYH2 gene may be an important driver of CD4+ T cell and macrophage infiltration.

In summary, the core genes MYL1, MYH7, MYL2, and MYH2 obtained from the analysis in this study are closely related to ES. Among them, MYL1 and MYL2 may be the main cause of the highly invasive nature of ES, and MYH7 and MYH2 are mainly involved in immune infiltration and play a role in the development of ES. The core genes analyzed in this study provide new ideas for the study of the highly invasive and metastatic biological characteristics of ES as well as drug therapy, and may provide new targets for ES drug development. However, this study was limited to bioinformatics analysis only and lacked the support of animal experiments or clinical data. The conclusions obtained in this study need to be further verified by basic experiments.

**Fund Projects**
National Natural Science Foundation of China (82260887); Guangxi Natural Science Foundation (2020GXNSFAA259068); The first batch of high-level talent research projects of the hospital affiliated with the Right River Ethnic Medical College (R20196321).

**Data Availability**
All data included or relevant to the study are available upon request by contact with the corresponding author.

**Authors’ Contributions**
Luchang CHEN, Huifang ZENG and Wujia YANG conceptualized and designed this study. Changtai LUO, Dong LUO, Zhenjing SI and Wei WANG analyzed and wrote the paper in the study. Haidong ZHOU finished the drawing and manuscript revision work.

**Acknowledgements**
This work was supported by the National Natural Science Foundation of China (82260887), Guangxi Natural Science Foundation (2020GXNSFAA259068), The first batch of high-level talent research projects of the hospital affiliated to the Right River Ethnic Medical College (R20196321), respectively.

**Conflicts of Interest**
All authors declare that there has not been any commercial or associative interest that represents competing interests in connection with the work submitted.

**References**


[23] Li, C., Guan, R., Li, W., Wei, D., Cao, S., Chang, F., et al. (2023) Analysis of Myosin Genes in HNSCC and Identify MYL1 as a Specific Poor Prognostic Biomarker, Promotes Tumor Metastasis and Correlates with Tumor Immune Infiltration in HNSCC. BMC Cancer, 23, Article No. 840. https://doi.org/10.1186/s12885-023-11349-5


[38] Ju, G., Yao, Z., Zhao, Y., Zhao, X. and Liu, F. (2023) Data Mining on Identifying Diagnosis and Prognosis Biomarkers in Head and Neck Squamous Carcinoma. Scientific Reports, 13, Article No. 10020. https://doi.org/10.1038/s41598-023-37216-8