Urinary and Renal Research (2022) Volume 3 Issue 1, *pp.*1–8.
doi:10.54517/urr.v3i1.2055

Bioinformatics analysis of the muscle-invasive bladder cancer subtypes

XU Wenbin, ZHENG Weiying, XIA Hong, HUA Lin
School of Biomedical Engineering, Capital Medical University, Beijing 100069
Corresponding author: HUA Lin(E-mail: hualin7750@ 139. Com)

Abstract: **Objective** In order to improve the accuracy in distinguishing subtypes of bladder cancer and to explore its potential therapeutic targets, we identify differences between two kinds of bladder cancer subtypes (basal-like and luminal) in molecular mechanism and molecular characteristics based on the bioinformatics analysis. **Methods** In this study, the RMA (robust multichip averaging) was applied to normalize the mRNA profile which included 22 samples from basal-like subtype and 132 from luminal subtype, and the differential expression analysis of genes with top 1000 highest standard deviation was performed. Then, the Gene Ontology and KEGG pathway enrichment analysis of differentially expressed genes was performed. In addition, the protein-protein interactions networks analysis for the top 100 most significant differentially expressed genes was performed. Results A total of 742 differentially expressed genes distinguishing basal-like and luminal subtypes were found, of which 405 were up-regulated and 337 genes were down-regulated in basal-like subtype. GO enrichment analysis showed that differentially expressed genes were significantly enriched in the extracellular matrix, chemotaxis and inflammatory response. KEGG pathway enrichment analysis showed that the differentially expressed genes were significantly enriched in the pathway of extracellular matrix receptor interaction. The hub proteins we founded in protein-protein interaction networks were LNX1, MSN and PPARG. Conclusions In this study, the mainly difference of molecular mechanism between basal-like and luminal subtypes are alteration in extracellular matrix region, cell chemotaxis and inflammatory response. Genes such as LNX1, MSN and PPARG were forecast to play important roles in the classification of bladder carcinoma subtypes.

Keywords: Bladder carcinoma; Molecular subtype; Differentially expressed gene; Enrichment analysis; Protein-protein interaction network

0 Introduction

Bladder cancer is a complex malignant tumor with multiple pathogenic factors. Its incidence rate ranks ninth among malignant tumors. The incidence rate of bladder cancer in men is three times that in women, and the incidence rate increases with age. Bladder cancer can be divided into urothelial carcinoma, squamous cell carcinoma and adenocarcinoma, of which 75%~80% are urothelial carcinoma [1]. Clinically, bladder cancer can be divided into

Fund Project: supported by Beijing Natural Science Foundation (7142015)

Author unit: School of Biomedical Engineering, Capital Medical University (Beijing 100069)
Corresponding author: Hua Lin. E-mail: hualin7750@ 139. Com

superficial bladder (Ta-T1), carcinoma in situ (TIS) and myometrial invasive bladder cancer (T2-T4) based on the degree of penetration or infiltration of primary tumors into the bladder wall [2]. Myometrial invasive bladder cancer has rapid progression, high malignancy, easy metastasis, recurrence and high mortality. Although there are many treatment methods, the prognosis is poor and the effect is not ideal. The 6-year survival rate accounts for about 25%[2].

The pathogenic factors of bladder cancer can be summarized as follows: chemical environment exposure, chronic stimulation and related gene abnormalities [3]. Among them, aromatic amines, aniline dyes, nitrite and nitrate, acrolein and arsenic are exposed in the chemical environment, but the most important environmental factor is smoking. Chronic irritation includes indwelling catheter, infection and pelvic radiotherapy. Gene abnormalities include variation of FGFR3[4], TP53[5], TSC1[5], amplification of PPARG[6], CCND1[6], EGFR[6] and deletion of CDKN2A[6], RB1[6].

Genome wide analysis showed that the main RNA expression subtypes of myometrial invasive bladder cancer were basal like and luminal [7], and different cancer subtypes had different biomarkers, which prompted the development of new treatment and auxiliary diagnostic tools. Compared with the lumen type, basal like bladder cancer has higher malignant degree and metastasis rate, poor prognosis and short survival [8]. Although basal bladder cancer is more invasive, it is highly sensitive to cisplatin combined chemotherapy. However, p53 like tumors in the lumen bladder cancer subtype are resistant to chemotherapy to a certain extent, which may have some problems for patients who need treatment [8]. By comparing the genes of the two subtypes of bladder cancer, it is found that EGFR[8], ZEB2[7,9] and other genes are overexpressed in basal like bladder cancer, and PPARG[7], FGFR3[4,7,9] and other genes are overexpressed in lumen bladder cancer. These genes are helpful to the study of cancer typing and the development of new treatment strategies and molecular targets for specific types of cancer. In this study, the gene expression profile data of two myometrial invasive bladder cancer subtypes, basal like and lumen bladder cancer, were used to screen the differentially expressed genes, annotate their functions, analyze the enrichment of pathways, and construct protein-protein interaction networks to further study the differences in the molecular mechanisms of the two bladder cancer subtypes.

1 Materials and methods

1.1 Data source and data preprocessing

The data source in this study is the gene expression profile data of bladder cancer samples in the French national CIT (cartes d'identite'des tumours) program. The data number is e-mtab-1940. From the mineica package of Bioconductor (http://www. Bioconductor. Org/packages/release/bioc/html /mineica. HTML). The data set composed of two bladder cancer subtypes included 22 basal like bladder cancer samples and 132 luminal bladder cancer samples. The robust multichip averaging (RMA) algorithm was used to standardize the data, and the first 1000 genes with the highest standard deviation were selected for analysis.

1.2 Screening of differentially expressed genes

BRB arraytools is an integrated data package for visualization and statistical analysis of microarray gene expression. In this study, BRB arraytools version 4.5.1 was used to identify the differentially expressed genes of

two bladder cancer subtypes. The method used was two sample t-test, and the test standard was $p<0.05$.

1.3 Enrichment analysis of differentially expressed genes

In order to further identify the differences between the two subtypes of myometrial invasive bladder cancer at the level of functional genomics, David (the database for annotation, visualization and integrated discovery) online analysis tool (https:// david. Ncifcrf. Gov /) perform go (gene ontology) functional annotation (including biological processes, molecular functions and cell components) and KEGG (Kyoto Encyclopedia of genes and genes) enrichment analysis on the two subtypes of differentially expressed genes. The screening thresholds were $p < 0.001$ and $fdr<0.05$.

1.4 Construction of protein-protein interaction network and recognition of hub protein

Current studies have found that systematic analysis of the interaction between a large number of proteins in biological systems is of great significance for understanding the working principle of proteins in biological systems, understanding the reaction mechanism of biological signals and energy metabolism under special physiological conditions such as diseases, and understanding the functional relationship between proteins. In this study, the top 100 genes with the most significant differences were selected and analyzed in the human protein reference database (HPRD, http://_www. Hprd. Org /) extract the interaction pairs containing the target genes, and then construct the protein interaction network with Cytoscape 3.4.0 Bioinformation Analysis software.

2 Results

2.1 Identification of differentially expressed genes

According to the test standard of $p< 0.05$, 742 differentially expressed genes were screened out from 1000 genes, among which 405 genes were up-regulated and 337 genes were down regulated in the basal like type compared with the lumen type. The gene heat map is shown in Figure 1.

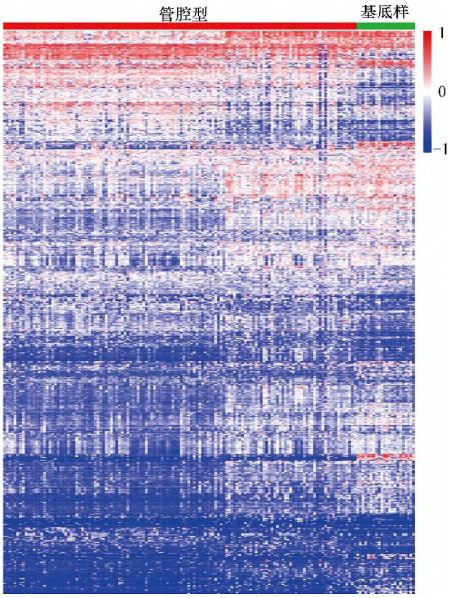


Figure 1 Heatmap of genes

2.2 Enrichment analysis of differentially expressed genes

The results of go enrichment analysis of differentially expressed genes are shown in Figure 2. Each column in the figure represents the GO term. The column length represents the number of genes enriched in this term (the higher the number of genes enriched in this term), and the column color represents the p value (from red to blue, the smaller the p value). In the study, the differentially expressed genes were most significantly enriched in the life

processes related to extracellular matrix, including the assembly and arrangement of extracellular matrix components and the decomposition of extracellular matrix. A large part is concentrated in the molecular

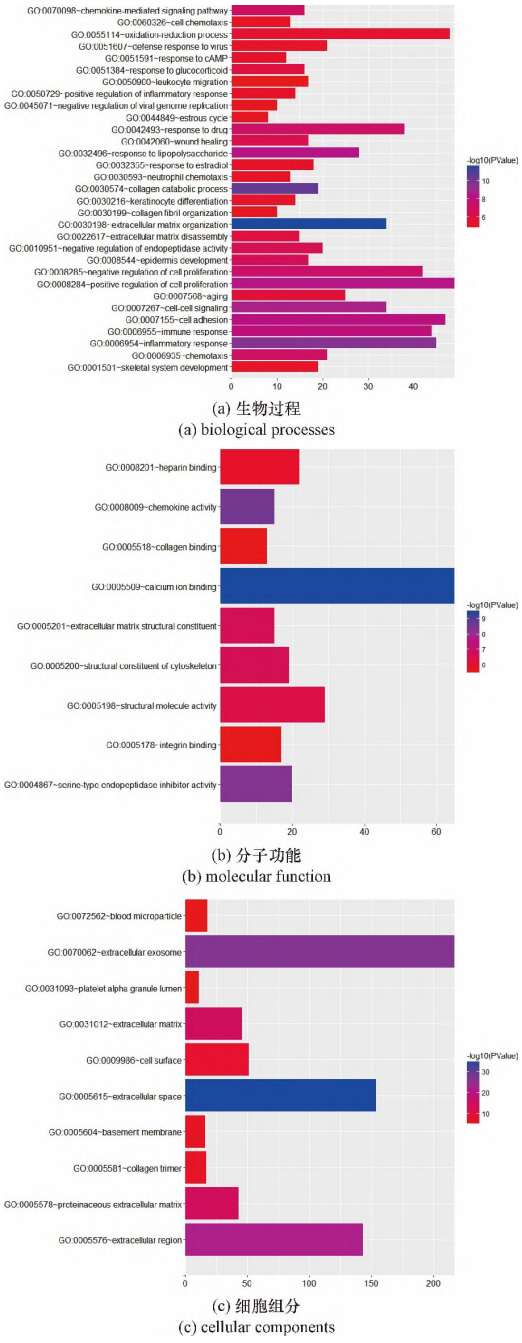


Figure 2 Go enrichment analysis

mechanisms related to cell chemotaxis, such as cell chemotaxis, neutrophil chemotaxis, chemokine mediated signal pathways and chemokine activity. In addition, the differentially expressed genes were also significantly enriched in the response to anti-inflammatory and anti-toxic glucocorticoids and cyclic adenosine monophosphate (camp). Camp can regulate the synthesis of neurotransmitters and promote the secretion of hormones. It is the second messenger in cells. Camp also regulates cell proliferation, differentiation and tissue development. Camp induces differentiation and inhibits the malignant phenotype of bladder cancer cells by down regulating the expression of Ras and myc genes, which plays an important role in the prevention and treatment of tumors [11].

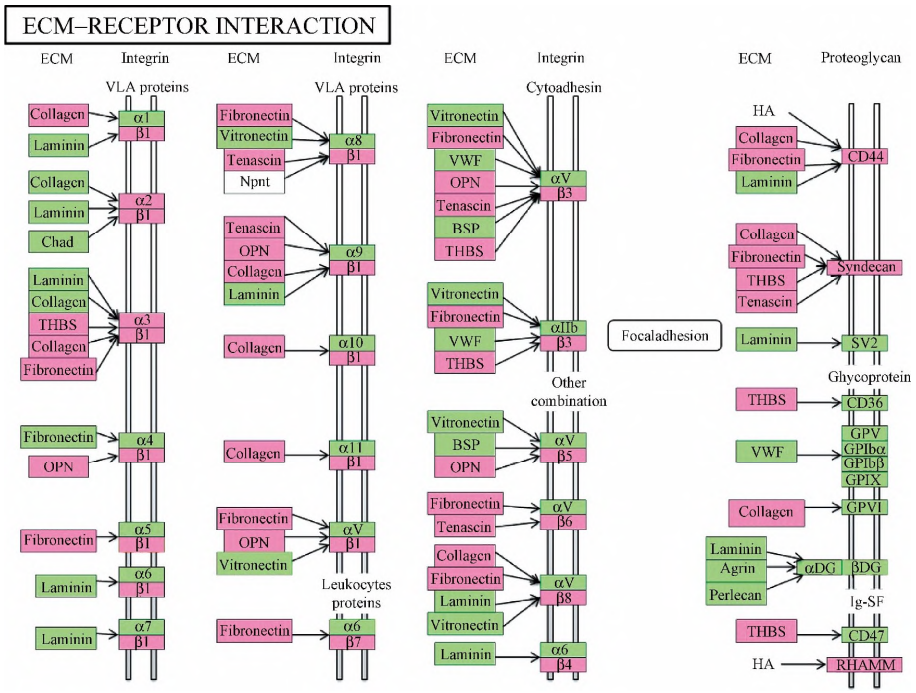
The KEGG pathway of differentially expressed genes was enriched by using David online analysis tool. The enrichment pathways are shown in Table 1. It can be seen that the most significant enrichment pathway of differentially expressed genes is the extracellular matrix (ECM) receptor interaction. In addition, it is also significantly enriched in adhesive plaque, complement and coagulation cascade, infection (including Staphylococcus aureus infection and amoebiasis). The extracellular matrix receptor interaction pathway includes 18 differentially expressed genes, as shown in Figure 3.

Table 1 KEGG pathway names and P values of differentially expressed genes

KEGG channel name	P
Hsa04512: ECM-receptor interaction	2.79E-07
Hsa04610: Complement and coagulation cascades	3.31E-07
Hsa05150: Staphylococcus aureus infection	3.99E-06
Hsa05146: Amoebiasis	4.99E-06
Hsa05323 :Rheumatoid arthritis	8.43E-06
Hsa04510:Focal adhesion	1.92E-05
Hsa00980:Metabolism of xenobiotics by cytochrome P450	2.42E-05

2.3 Construction of protein-protein interaction network and recognition of hub protein

The selected differentially expressed genes were mapped to the human protein interaction database HPRD and a protein interaction network was constructed. The protein-protein interaction network identified 355 genes and 332 interactions. In protein-protein interaction networks, the nodes with high connectivity are called hub, which are important genes or their translation products supporting the basic activities of life. At the same time, the central node is significantly enriched with genes related to genetic diseases such as cancer.



The pink box indicates the differentially expressed genes; the green box indicates non differentially expressed genes

Figure 3 Metabolic pathway of extracellular matrix receptor interaction

In the constructed protein-protein interaction network, the central nodes are lnx1 (connectivity 47), PPARG (connectivity 33) and MSN (connectivity 20). Therefore, it can be considered that these genes play an important role in the subtype of bladder cancer. Lnx1 gene is one of the four members of LNX ubiquitin ligase family. Its structure includes N-terminal ring domain and 4 C-terminal PDZ domains. This gene participates in notch, neuregulin-1/erbb and other important signal pathways, participates in the recombination of cell tight junctions, and plays an important role in tumorigenesis [12]. MSN (moesin) is a member of ERM family, including Ezrin and Radixin. ERM protein acts as a cross-linking agent between plasma membrane and actin based cytoskeleton. Moesin is located in filopodia and membrane processes, and plays an important role in intercellular recognition, signal transduction and cell movement [13]. PPARG gene encodes a member of peroxisome proliferator activated receptor (PPAR) subgroup-γ Protein [14]. PPARG can regulate cell growth, immune surveillance and adipocyte differentiation [14]. In the urinary system,

PPARG activation induces the changes of urothelial cells and eventually leads to the expression of specific markers of terminal urothelial differentiation [14-15].

3 Discussion

Bladder cancer is a heterogeneous disease characterized by complex molecular changes and abnormal gene expression. Myometrial invasive bladder cancer is often associated with cancer metastasis, which leads to high mortality. Genome wide mrna expression profiling has been widely used to study the molecular heterogeneity of human bladder cancer and to explore the gene expression characteristics related to cancer progression, metastasis and survival. The basal and lumen subtypes of myometrial invasive bladder cancer have different clinical features, molecular characteristics and molecular targets. Therefore, through bioinformatics analysis, it is of great significance to reveal the molecular mechanism of the occurrence and development of the two subtypes of bladder cancer at the molecular level, so as to improve the level of diagnosis and treatment of bladder cancer.

The emergence of high-throughput data analysis technology has accelerated the development of cancer research. Recent genomic studies have found a large number of biomarkers for basal like and bureaucratic bladder cancer. Choi team [8] found Δ Np63 α The gene was significantly enriched in basal like bladder cancer and controlled the expression of biomarkers in basal like bladder cancer. In basal like bladder cancer, Δ Np63 α It controls the adhesion between epithelial cells and extracellular matrix in basal like tumors, which is similar to the results of this study. In the go enrichment analysis, it was found that the differentially expressed genes of the two bladder cancer subtypes were most significantly enriched in the life activities related to the extracellular space, and through the KEGG pathway analysis, it was found that the differentially expressed genes were also significantly enriched in the extracellular matrix receptor interaction pathway. The occurrence, development, invasion and metastasis of malignant tumors are often accompanied by changes in the expression of extracellular matrix and its cell surface receptors [16]. Specific interactions between cells and extracellular matrix can directly or indirectly control cell activity, such as adhesion, migration, differentiation, proliferation and apoptosis. Metastatic tumor cells interact with matrix in many stages of tumor deterioration and metastasis. Some specific types of stroma and basement membrane play a key role in the progression of malignant tumors and blood borne dissemination [16].

Since the discovery of HRAS as the first oncogene in bladder cancer cell line, many genes mutated in bladder cancer have been identified, such as TP53, Rb1, TSC1, FGFR3 and PIK3CA.

The researchers found that there was a significant correlation between the classification of bladder cancer subtypes and gene mutations [5]. FGFR3 and TSC1 mutations frequently occur in lumen subtypes, while RB1 pathway changes including RB1 mutation/deletion, CCND1 amplification, SOX4/E2F3 amplification and CCNE1 amplification frequently occur in basal subtype bladder cancer [6-7]. The hub proteins in the constructed protein interaction network are lnx1, MSN and PPARG. It can be considered that the genes encoding these proteins play an important role in the typing of myometrial invasive bladder cancer. PPARG is the most well-known biomarker

among the three genes in the field of bladder cancer subtype research. Many studies have reported on this gene. In the copy number detection of bladder cancer, it can be found that PPARG gene is significantly amplified in lumen type bladder cancer [5, 8, 10], so the overexpression of PPARG gene has become an important marker of lumen type bladder cancer. PPARG is an oncogene related to differentiation in bladder cancer, and PPARG agonists can inhibit the growth of bladder cancer cell lines in vivo and in vitro [17], which indicates that PPARG will be an important target for the treatment of bladder cancer once the molecular subtype specific effects of cancer cell biology are better understood. At present, there are few reports on the role of lnx1 and MSN in the subtype of bladder cancer, but the variation of these two genes plays an important role in the carcinogenesis of bladder cancer. So far, there have been few studies on molecular subtypes of bladder cancer. Lnx1 and MSN in this study may be a new discovery, and may also become new therapeutic targets, which requires a large number of clinical trials and analysis to verify.

Recent studies have found that bladder cancer and breast cancer share some genetic similarities. Through genome-wide mrna expression profile analysis, it is found that breast cancer can be roughly divided into five subtypes: claudin low, basal like, luminal a, luminal B and HER2 enriched. Among them, breast cancer has low density, basal like subtype and bladder cancer's basal like subtype. Breast cancer's luminal a subtype is highly similar to bladder cancer's luminal subtype, It has similar survival time, prognosis and genetic signal pathway [9]. It is worth mentioning that the MSN gene found in this study is also differentially expressed in the basal like and lumen subtypes of breast cancer [13], which further illustrates the genetic similarity between bladder cancer and breast cancer. If bladder cancer subtype can play a role in the treatment classification like breast cancer subtype, it will be particularly helpful for doctors to develop the best treatment process for patients, which is also one of the tasks to be done in the future.

4 Conclusion

In this study, bioinformatics methods were used to analyze the mrna expression profile data of two bladder cancer subtypes, and the differences in biological functions and signal pathways between the two bladder cancer subtypes were found, and some key genes lnx1, MSN and PPARG that may help to distinguish bladder cancer subtypes were found. However, due to the limited data set of bladder cancer subtypes, more data set analysis is needed to ensure the stability of the research results. These results can be used as the basis for the project design, and as a clue to further explore the molecular mechanism of bladder cancer subtypes.

References

[1] Bryan RT. Update on bladder cancer diagnosis and management[J]. Trends in Urology & Mens Health, 2013, 4(5) : 7–11.

[2] Mao JW, Li QW. Progress in the treatment of muscular invasive bladder cancer[J]. Chinese Journal of General Practice, 2013, 11 (5) : 778-780.

[3] Kaufman DS, Shipley WU, Feldman AS. Bladder cancer [J]. Lancet, 2009, 374(374) : 239-249.

[4] Xing AI, Jia Z, Wang J, et al. Bioinformatics analysis of the target gene of fibroblast growth factor receptor 3 in bladder cancer and associated molecular mechanisms[J]. Oncology Letters, 2015,

10 (1) : 543-549.

[5] Martin-Doyle W, Kwiatkowski DJ. Molecular biology of bladder cancer[J]. Hematology/oncology Clinics of North America, 2015, 29(2) : 5-12.

[6] The Cancer Genome Atlas Research Network. Comprehensive Molecular Characterization of Urothelial Bladder Carcinoma[J]. Nature, 2014, 507(7492) : 315-322.

[7] Choi W, Porten S, Kim S, et al. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy[J]. Cancer Cell, 2014, 25(2) : 152-65.

[8] Choi W, Czerniak B, Ochoa A, et al. Intrinsic basal and luminal subtypes of muscle-invasive bladder cancer[J]. Nature Reviews Urology, 2014, 11(7) : 400-410.

[9] Damrauer JS, Hoadley KA, Chism DD, et al. Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology[J]. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111(8) : 3110-3115.

[10] Biton A, Bernardpierrot I, Lou Y, et al. Independent component analysis uncovers the landscape of the bladder tumor transcriptome and reveals insights into luminal and basal subtypes [J]. Cell Reports, 2014, 9(9) : 1235-1245.

[11] Hu CX, Zhang JR, Huang LH, et al. Application of cyclic adenosine monophosphate in clinical application of tumor[J]. Tianjin Pharmacy, 2013, 25(6) : 49-52.

[12] Guo ZG, Gao YH. Expression, activity determination of ubiquitin ligase LNX1 and primary study on its function[J]. Letters in Biotechnology, 2012, 23(4) : 523-526.

[13] Wang CC, Liau JY, Lu YS, et al. Differential expression of moesin in breast cancers and its implication in epithelial-mesenchymal transition[J]. Histopathology, 2012, 61(61) : 78-87.

[14] Conconi D, Panzeri E, Redaelli S, et al. DNA copy number alterations and PPARG amplification in a patient with multifocal bladder urothelial carcinoma[J]. BMC Research Notes, 2012, 5 (1) : 607.

[15] Mansure J J, Nassim R, Chevalier S, et al. A novel mechanism of PPAR gamma induction via EGFR signalling constitutes rational for combination therapy in bladder cancer[J]. Plos ONE, 2013, 8 (2) : e55997.

[16] Li ZY. Extracellular matrix and bladder tumor metastasis[J]. International Journal of Urology and Nephrology, 2012, 32 (3) : 363-367.

[17] Yan S, Yang X, Chen T, et al. The PPAR γ agonist troglitazone induces autophagy, apoptosis and necroptosis in bladder cancer cells[J]. Cancer Gene Therapy, 2014, 21(5) : 188-193.