RESEARCH ARTICLE

Screening of Differentially Expressed Genes Related to Bladder Cancer and Functional Analysis with DNA Microarray

Yi-Dong Huang*, Wei Shan, Li Zeng, Yang Wu

Abstract

Objective: The purpose of this study was to identify genes related to bladder cancer with samples from normal and disease cases by microarray chip. Methods: After downloading the gene expression profile GSE3167 from Gene Expression Omnibus database which includes 50 bladder samples, comprising 9 normal and 41 disease samples, differentially expressed genes were identified with packages in R language. The selected differentially expressed genes were further analyzed using bioinformatics methods. Firstly, molecular functions, biological processes and cell component analysis were researched by software Gestalt. Then, software String was used to search interaction relationships among differentially expressed genes, and hub genes of the network were selected. Finally, by using plugins of software Cytoscape, Mcode and Bingo, module analysis of hub-genes was performed. Results: A total of 221 genes were identified as differentially expressed by comparing normal and disease bladder samples, and a network as well as the hub gene C1QBP was obtained from the network. The C1QBP module had the closest relationship to production of molecular mediators involved in inflammatory responses. Conclusion: We obtained differentially expressed genes of bladder cancer by microarray, and both PRDX2 and YWHAZ in the module with hub gene C1QBP were most significantly related to production of molecular mediators involved in inflammatory responses. From knowledge of inflammatory responses and cancer, our results showed that, the hub gene and its module could induce inflammation in bladder cancer. These related genes are candidate bio-markers for bladder cancer diagnosis and might be helpful in designing novel therapies.

Keywords: Bladder cancer - differentially expressed gene - interaction network - function enrichment analysis - module analysis

Asian Pac J Cancer Prev, 14 (8), 4553-4557

Introduction

Bladder cancer (BC) is estimated to be the ninth most common cause of cancer worldwide and the 13th most numerous cause of death from cancer (Parkin, 2008). It remains the fifth leading new cancer diagnosis in the United States (US), the fourth most common malignancy in men and the ninth most common in women in the US (Siegel et al., 2008). Healthcare costs for BC treatment are among the highest of any malignancy (Avritscher et al., 2006). More than 90% of urinary bladder cancers (UBC) present as transitional cell (urothelial) carcinoma, while approximately 5% are squamous cell carcinoma, less than 2% are adenocarcinoma, and the remainder are other rare subtypes (Heney, 1992). Modern cigarettes are proposed as a possible mechanism for the increased risk of BC (Baris et al., 2009).

Over the last decade, scientists have demonstrated multiple interacting molecular pathways are present in the genesis of BC. There are more alterations in the receptor tyrosine kinase–Ras pathway, especially the mutation of HA-RAS (HRAS) and fibroblast growth factor receptor 3 (FGFR3) genes in low-grade non-invasive BC than normal tissue (Pasin et al., 2008). Alterations in the p53, phosphatase and tensin homolog (PTEN), retinoblastoma (Rb) pathways are more prevalent in high-grade non-invasive and invasive tumors (Mitra et al., 2006; Puzio-Kuter et al., 2009). Numerous factors within the mTOR-signaling cascade have also appear to be altered in BC (Chingand Hansel, 2010).

By using high-throughput tissue microarray analysis, 11q13 gene is found amplification in BC (Zaharieva et al., 2003). However, there are not many high-throughput screening researches and function identification of differential gene expression about it. In this study we performed an analysis of genomic microarray based on BC expression data and got a hub gene C1QBP, which, together with both PRDX2 and YWHAZ in its modules, are involved in inflammatory response. Our results show that, the hub gene and its modules might play important roles on the conversion from inflammation to carcinoma in bladder.

Materials and Methods

Affymetrix microarray data

The gene chip GSE3167 which downloaded from the gene expression database Gene Expression Omnibus...
(GEO) (Dyrskjot et al., 2004) was a sixty data peripheral blood samples of bladder cancer formed by nine normal samples and forty-one bladder cancer samples. The remaining ten samples have nothing to do with the analysis, so they are not selected. The platform Information is that: GPL96 (HG-U133A), with Affymetrix Human Genome U133A Array. Annotation information of chip probe was from Affymetrix, which contains all Affymetrix ATH1 (25K) chip probe, all the original file and the platform probe annotation information file were download.

Data preprocessing and gene differences analysis

By using R language Affy package, the original data was preprocessed (Troyanskaya et al., 2001; Fujita et al., 2006). The limma in R language package was used to analyze for differential gene expression in all the nine normal samples and forty-one bladder cancer chip. Bayesian methods was performed to do multiple comparison correction (Benjaminiand Hochberg, 1995). Threshold value for the most differentially expressed genes were set as \( p \)-value < 0.01 and |logFC| > 1.

Enrichment analysis of differentially expressed genes

Gene Set Analysis Toolkit (Gestalt) tool was used to do enrichment analysis on the differentially expressed genes in three sections of function, biological process and molecular composition. Gestalt is a suite rich of analysis of biologically relevant content collecting eight species, including human, rat, mouse, and other data from various different public data resources, such as NCBI, Ensemble, Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Gene Ontology (GO) (Wen et al., 2011). Differentially expressed genes were screened by using enrichment analysis based on the hypergeometric distribution algorithm of Gestalt.

Constructing protein interaction network

Protein-protein interactions (PPIs) had been marked as the main actors for all of the processes taking place in a cell and therefore great efforts had been made towards the understanding of their biological function (Tong et al., 2002). By using the online database resource Search Tool for the Retrieval of Interacting Genes (STRING) (Szklarczyk et al., 2011), interaction between the differentially expressed genes were predicted. Based on the connection of major database, string database scored gene interactions from a variety of angles, such as homology, experiments and text mining (Tong et al., 2002).

Hub protein screening

Studies on biology network indicated that majority of biological networks obey scale-free (scale-free) network attributes (Jeong et al., 2000), which mean a small number of nodes in the network had a large number of connections, most nodes had only a few connections, these few nodes were the network critical node (hub) (Yu et al., 2004). By using the interaction network node and the scale-free nature of interaction in the protein network, the statistics node and its distribution were analyzed and the center of the net, the hub protein was found.

Interaction network functional modules

Proteins in a PPI network and the same module usually complete the same biological processes and functions by co-expression. In this study, after got network visualization by cytoscape (Duncan et al., 2010), plugin Mcode (Goldberg et al., 2007) module parameters: Degree cutoff ≥ 2, K-core ≥ 2) were used to divide the network module and Bingo (Dunham et al., 2012) was used to annotate the function of hub gene functions based on the hypergeometric distribution (adj \( p \)-value < 0.01).

Results

Table 1. Enriching Sequence of Differentially Expressed Genes

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Func</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOTERM_MF_FAT</td>
<td>GO:0051082</td>
<td>unfolded protein binding</td>
<td>2.33E-08</td>
</tr>
<tr>
<td></td>
<td>GO:0003723</td>
<td>RNA binding</td>
<td>0.000154</td>
</tr>
<tr>
<td></td>
<td>GO:0001666</td>
<td>nucleotide binding</td>
<td>0.042638</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0000398</td>
<td>nuclear mRNA splicing, via spliceosome</td>
<td>0.040382</td>
</tr>
<tr>
<td></td>
<td>GO:0003777</td>
<td>RNA splicing</td>
<td>0.040382</td>
</tr>
<tr>
<td></td>
<td>GO:0006457</td>
<td>protein folding</td>
<td>0.0000112</td>
</tr>
<tr>
<td></td>
<td>GO:0008380</td>
<td>RNA splicing</td>
<td>0.000191</td>
</tr>
<tr>
<td></td>
<td>GO:0066397</td>
<td>mRNA processing</td>
<td>0.0000393</td>
</tr>
<tr>
<td></td>
<td>GO:0016071</td>
<td>mRNA metabolic process</td>
<td>0.00000258</td>
</tr>
<tr>
<td></td>
<td>GO:0066396</td>
<td>RNA processing</td>
<td>0.00254</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0042470</td>
<td>melanosome</td>
<td>6.55E-12</td>
</tr>
<tr>
<td></td>
<td>GO:0048770</td>
<td>pigment granule</td>
<td>6.55E-12</td>
</tr>
<tr>
<td></td>
<td>GO:0030485</td>
<td>smooth muscle contractile fiber</td>
<td>0.008425</td>
</tr>
<tr>
<td></td>
<td>GO:0016023</td>
<td>cytoplasmic membrane-bounded vesicle</td>
<td>0.000141</td>
</tr>
<tr>
<td></td>
<td>GO:0031988</td>
<td>membrane-bounded vesicle</td>
<td>0.000251</td>
</tr>
<tr>
<td></td>
<td>GO:0031410</td>
<td>cytoplasmic vesicle</td>
<td>0.00213</td>
</tr>
<tr>
<td></td>
<td>GO:0031982</td>
<td>vesicle</td>
<td>0.004381</td>
</tr>
<tr>
<td></td>
<td>GO:0070013</td>
<td>intracellular organelle lumen</td>
<td>0.0000135</td>
</tr>
<tr>
<td></td>
<td>GO:0043233</td>
<td>organelle lumen</td>
<td>0.00000308</td>
</tr>
<tr>
<td></td>
<td>GO:0031974</td>
<td>membrane-enclosed lumen</td>
<td>0.00000603</td>
</tr>
<tr>
<td></td>
<td>GO:0043228</td>
<td>non-membrane-bounded organelle</td>
<td>0.007334</td>
</tr>
<tr>
<td></td>
<td>GO:0043232</td>
<td>intracellular non-membrane-bounded organelle</td>
<td>0.007334</td>
</tr>
</tbody>
</table>
Table 2. Annotation List for Function of the Module

<table>
<thead>
<tr>
<th>GO-ID</th>
<th>p-value</th>
<th>corr p-value</th>
<th>Description</th>
<th>Genes in test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>2532</td>
<td>0.0000114</td>
<td>0.0045</td>
<td>production of molecular mediator involved in inflammatory response</td>
<td>YWHAZ/PRDX2</td>
</tr>
<tr>
<td>6986</td>
<td>0.00158</td>
<td>0.0421</td>
<td>response to unfolded protein</td>
<td>HSP90AB1/HSPA5</td>
</tr>
<tr>
<td>45454</td>
<td>0.00158</td>
<td>0.0421</td>
<td>cell redox homeostasis</td>
<td>PRDX4/PRDX2</td>
</tr>
<tr>
<td>6457</td>
<td>0.000432</td>
<td>0.0421</td>
<td>protein folding</td>
<td>HSP90AB1/PFDN2/PPIB</td>
</tr>
<tr>
<td>6916</td>
<td>0.000684</td>
<td>0.0421</td>
<td>anti-apoptosis</td>
<td>YWHAZ/PRDX2/HSPA5</td>
</tr>
</tbody>
</table>

221 genes which met the difference threshold ($p < 0.01$ and $|\log FC| > 1$) were screened, in which included 40 down-regulated and 181 up-regulated genes (Table 1).

Enrichment analysis of differential expression genes

Enrichment analysis of function, biological process and molecular composition were carried out by Gestalt. Significant enrichment clusters of function, biological process and molecular composition were 3, 7 and 12, as shown in Table 1 and Figure 2. As shown in column height of the Figure 2, among all the differentially expressed genes, most enriched in the above three aspects of the genes were: RNA binding function, RNA processing process and the non-membrane organelles intracellular.

Screening protein interaction network of differentially expressed genes

Differentially expressed gene product protein interaction was searched by String Software. A total of 801 pairs (189) differences were got in gene interaction network. Protein interaction network of protein interactions between coefficients is shown in Table 2.

Screening for hub protein

Calculated node degree of the differentially expressed gene product which previously obtained from the constituted interaction networks, the highest node was the gene C1QBP (node 40) as shown in Figure 3, which mean this gene has an important role in the network.

Analyzing of module functions in co-expression interaction network

By using Mcode which is the plugin of Cytoscape to partition off the network module, module contained the hub gene was found, then garnished with another plugin Bingo to division and comments the whole network function module based on the hypergeometric distribution (adj $p < 0.01$). A module that contained the hub gene C1QBP was obtained (Figure 3), and another 5 notable features were got from this module (Table 2). Inflammatory response functions which contained PRDX2 and YWHAZ are involved in cell proliferation was most significant in this module.

Figure 1. Normalized Expression Value Data Box Plots.
The blue boxes show nine normal bladder sample and pink ones show fort-one bladder cancer samples. The box in the black line means the median of each set of data, which determine the degree of standardization of data through its distribution, black line in the box is almost in the same straight line, indicating good degree of standardization.

Figure 2. Enrichment Pathways of Differentially Expressed Genes in Three Aspects. From left to right were the biological processes, cellular components and molecular functions. Abscissa axis represents the enriched function, the vertical axis represents the number of genes and the height of the pillars represents the number of genes involved in the function.

Figure 3. Differentially Expressed Gene Co-expression Network. Hub genes are shown in diamond-shaped node. The circular node indicates the other differentially expressed genes constituted interaction networks, the highest node was the gene C1QBP (node 40) as shown in Figure 3, which mean this gene has an important role in the network.
Discussion

Molecular markers which have the potential to overcome the limitations of conventionally used prognostic risk factors are promising tools in BC. The detection of molecular pathways involved in the bladder carcinogenesis might also be therapeutically useful for providing an opportunity to diagnose tumours earlier, identify those patients most at risk of disease recurrence and predict tumours respond to targeted therapies (Proctor et al., 2010; Gakis et al., 2012). An increasing number of studies have focused on the gain of the predictive accuracy of biomarkers when combined with standard clinical and pathological parameters in the last decade. They are urinary fibronectin (FN), the nuclear matrix protein 22 (NMP22) and bladder tumor antigen (BTA) (Malmström, 1993). Although there were many researches about the molecular biomarker of BC, however, because most studies are retrospective, so their predictive value in terms of addressing the need for additional therapy in an individual patient remains unclear. In this article, we obtained differentially expressed feature genes of BC, in which the module with hub gene CIQBP was most significantly and was related to production of molecular mediator involved in inflammatory response.

CIQBP/p32, which is also represented as synonyms of HABP1 (Hyaluronan Binding Protein 1) is one of hyaladherin implicated in cellular functions like adhesion and tumor invasion (Guptaand Datta, 1991), sperm maturity and motility (Ranganathan et al., 1994) and cellular signaling (Rao et al., 1996). It presents in human chromosome 17p12-13 (accession number NP_001203) and encodes a sialic acid containing a pro-protein sequence of 282 amino acids, consist of the first 73 amino acids which are cleaved off during protein mature procedure and the rest mature protein of 209 amino acids (Deband Datta, 1996). CIQBP protein is a doughnut shaped trimer (Jiang et al., 1999) that is primarily localized in the mitochondria (Dedioand Mueller-Esterl, 1996; Matthewsand Russell, 1998), has also been reported to present at the cell surface (Gebrehiwet et al., 1994; Dedio et al., 1996) and in the nucleus (Brokstad et al., 2001; Majumdar et al., 2002). Human p32 has been reported to interact with human immunodeficiency virus Tat (Rubinstein et al., 2004), mediated the cellular apoptotic response (Kamaland Datta, 1998), and has a link to autophagy and cell cycle (Sengupta et al., 2005; Reef et al., 2007). Recent report also pointed that this p32 may be a novel marker of clinical progression in prostate cancer (Amamoto et al., 2011).

Peroxiredoxin 2 (PRDX2), mapped at 13q12, has both proliferative and antia apoptotic properties and thus may induce carcinogenic changes (Noh et al., 2001). When overexpressed, PRDX2 protects cancer cells from oxidative stress and thus mediates resistance to chemotherapeutic drugs and radiotherapy (Soini et al., 2006; Smith-Pearson et al., 2008). Recent research indicates that PRDX2 is involved in the negative regulation of inflammatory response in psoriatic skin (Ryu et al., 2011).

YWHAZ, as known as 14-3-3ζ, has been implicated in the initiation and progression of cancer and shown to be overexpressed in multiple cancer types (Neal and Yu, 2010). Several studies have suggested that YWHAZ expression contributes to the transformation phenotype. In breast cancer cell lines, overexpression of YWHAZ enhances anchorage-independent growth and inhibits stress induced apoptosis (Neal et al., 2010). In breast cancer, compared to normal ductal epithelium, YWHAZ is found improved in metastatic cells (Zang et al., 2004). Also in lung cancer, knockdown of YWHAZ sensitizes cells to stress-induced apoptosis and enhanced cell adhesion and cell-cell contacts (Niemantsverdriet et al., 2008). Similarly, knockdown of YWHAZ restores the sensitivity of lung cancer cells to anoikis and impaires their anchorage-independent growth (Li et al., 2008).

In the present study, we reported that the hub gene CIQBP in BC microarray is differentially expressed and in the module which CIQBP involved, PRDX2 is related to inflammatory response in psoriatic skin (Ryu et al., 2011), while YWHAZ is shown to be overexpressed in multiple cancer types (Neal et al., 2010). Although there are no articles show direct relationship among the here genes, our results discovered that both PRDX2 and YWHAZ were involved in the inflammatory response in the BC microarray chip samples. There is mounting evidence that inflammation plays critical roles in the pathogenesis of bladder cancer (Zhu et al., 2012). So we hypothesized that through the inflammatory response in the bladder tissue, the key gene CIQBP and its module have function from inflammation in bladder tissue to bladder cancer.

References


Screening and Functional Analysis of Differentially Expressed Genes Related to Bladder Cancer


