Bioinformatics analysis of differentially expressed proteins in prostate cancer based on proteomics data

Chen Chen,1 Li-Guo Zhang,1 Jian Liu,1 Hui Han,1 Ning Chen,1 An-Liang Yao,1 Shao-San Kang,1 Wei-Xing Gao,1 Hong Shen,2 Long-Jun Zhang,1 Ya-Peng Li,1 Feng-Hong Cao,1 Zhi-Guo Li3

1Department of Urology, North China University of Science and Technology Affiliated Hospital, 2Department of Modern Technology and Education Center, 3Department of Medical Research Center, International Science and Technology Cooperation Base of Geriatric Medicine, North China University of Science and Technology, Tangshan, People’s Republic of China

Abstract: We mined the literature for proteomics data to examine the occurrence and metastasis of prostate cancer (PCa) through a bioinformatics analysis. We divided the differentially expressed proteins (DEPs) into two groups: the group consisting of PCa and benign tissues (P&b) and the group presenting both high and low PCa metastatic tendencies (H&L). In the P&b group, we found 320 DEPs, 20 of which were reported more than three times, and DES was the most commonly reported. Among these DEPs, the expression levels of FGG, GSN, SERPINC1, TPM1, and TUBB4B have not yet been correlated with PCa. In the H&L group, we identified 353 DEPs, 13 of which were reported more than three times. Among these DEPs, MDH2 and MYH9 have not yet been correlated with PCa metastasis. We further confirmed that DES was differentially expressed between 30 cancer and 30 benign tissues. In addition, DEPs associated with protein transport, regulation of actin cytoskeleton, and the extracellular matrix (ECM)–receptor interaction pathway were prevalent in the H&L group and have not yet been studied in detail in this context. Proteins related to homeostasis, the wound-healing response, focal adhesions, and the complement and coagulation pathways were overrepresented in both groups. Our findings suggest that the repeatedly reported DEPs in the two groups may function as potential biomarkers for detecting PCa and predicting its aggressiveness. Furthermore, the implicated biological processes and signaling pathways may help elucidate the molecular mechanisms of PCa carcinogenesis and metastasis and provide new targets for clinical treatment.

Keywords: bioinformatics analysis, differentially expressed proteins, occurrence, literature mining, metastasis, prostate cancer, proteomics

Introduction

Prostate cancer (PCa) is one of the most common cancers in men, and it is the second leading cause of morbidity and mortality in men.1 The American Cancer Society predicts that 2,20,800 new cases of PCa will be diagnosed and 27,540 PCa-related deaths will occur in 2015.2 However, the pathogenesis and risk factors of PCa are complex and not yet fully understood. It has been well known that age, race, and heredity are risk factors of PCa. In addition, some exogenous factors, such as high-fat diet and less intake of vitamin E, selenium, lignans, or isoflavones would potentially promote the progression of PCa from latent to clinical type.3,4

With the development of proteomics technology, an increasing number of studies have examined PCa using proteomics, which is a powerful tool to investigate the pathogenesis and metastasis of PCa. Specifically, several tumor-related proteins have been identified in the tissue and biological fluids of various types and grades of cancer.5 For example, a previous global proteomics analysis, using the iTRAQ technology,
found that periostin was significantly upregulated in the PCa tissues compared with that in the benign prostatic hyperplasia (BPH). Periostin interacts with multiple cell-surface receptors, most notably integrins to promote cancer cell survival, epithelial–mesenchymal transition (EMT), invasion, and metastasis through signaling mainly via the PI3K/Akt and other pathways. Applying proteomics technology, Skvortsov et al discovered that lamin A was statistically highly discriminatory between low and high Gleason score for PCa and might serve as a new biomarker of tumor differentiation. These tumor-related proteins may offer clues of not only a promising biomarker for the prognosis of PCa but also a potential target for therapeutic interventions. The proteome more accurately reflects the dynamic state of a cell, tissue, or organism, and proteomics is anticipated to elucidate the mechanisms related to PCa. However, proteomics technologies require complex instrumentation and expensive consumables. Therefore, most studies have investigated small samples with great sampling error, which results in low reliability compared with the low-throughput methods (Western blot and enzyme-linked immunosorbent assay [ELISA]). In addition, current proteomics methods cannot resolve all proteins in one sample due to limitations in resolution. Different complementary methods can resolve proteins from the same sample; thus, a combination of different methods may resolve more proteins. For example, two-dimensional gel electrophoresis is associated with a low detection rate for low-abundance, hydrophobic, acidic, basic, small and large proteins, and the combination of liquid chromatographic techniques with a subsequent mass spectrometric analysis (LC-MS) is limited by the incomplete chromatographic separation of peptides. However, combining these two approaches will increase the detection rate. Moreover, proteomics studies have generated a large amount of data, but these data have not been thoroughly investigated. Consequently, data have accumulated, but their biological significance has not yet been determined. Proteomics experiments yield a list of identified proteins, and this list usually is not deeply investigated. Specifically, only the expression levels of three to five proteins are confirmed by Western blotting or immunofluorescence. Therefore, the biological significance of only a few proteins in the dataset is explored, and the remaining information, if deposited in a proteomic database, is lost or almost never investigated.

The application and development of computer technology and mathematics in the field of biology (bioinformatics) has become one of the most important tools in proteomics. Bioinformatics tools are essential for converting raw proteomics data into relevant knowledge and subsequently into useful applications. Furthermore, bioinformatics provides a method to convert datasets into biologically interpretable results and functional outcomes. Many studies have successfully combined data mining with bioinformatics technology. Hu et al analyzed breast cancer proteomics and genomics data using the literature-mining tool MedGene. They identified a set of relatively understudied yet highly expressed genes in estrogen receptor-negative tumors that warrant further examination. Li et al analyzed specific genes in androgen-independent PCa using literature mining and bioinformatics. They found that matrix metalloproteinase (MMP)-9 and epidermal growth factor receptor (EGFR) play important roles in the transformation of androgen-dependent PCa into androgen-independent PCa. These bioinformatics analyses have shed light on carcinogenesis at the molecular level. However, these studies usually rely on automatic data-mining tools, which only search the abstracts instead of the entire article. Therefore, differentially expressed genes or proteins that were reported in the articles may be overlooked. In addition, bioinformatics analyses of the whole proteome have not been employed to examine PCa occurrence and metastasis. Thus, in this study, we manually searched the proteomics literature of differentially expressed proteins (DEPs) related to PCa occurrence and metastasis in PubMed. Specifically, we searched for articles indexed before June 1, 2015 that were published in English. We hypothesized that proteins that had been repeatedly identified by multiple proteomics studies may serve as potential PCa biomarkers. Consequently, the frequency with which DEPs had been reported was counted, and frequently reported proteins were subjected to experimental verification. We also used bioinformatics tools to carry out a GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways analysis of DEPs to observe and analyze the changes in proteins and signaling pathways during PCa occurrence and metastasis from a global perspective.

**Materials and methods**

**The literature search strategy**

We used PubMed as the main source for the literature search. To prevent overlooking relevant articles, we first searched all proteomics studies related to PCa. We then excluded literature not directly related to human PCa occurrence and metastasis. We used the advanced search option in PubMed by inputting “(prostate cancer[Title]) AND Proteomics” and restricted the search to studies published in English before June 1, 2015. Studies that met the following criteria were excluded: 1) studies examining nonhuman tissue or cell lines,
2) studies examining the effects of certain interventions on the protein expression profiles, 3) studies that did not supply names or accession numbers of DEPs, and 4) literature reviews. The Ethics Committee of North China University of Science and Technology approved this study.

**Official gene symbol transform**

The DEPs were transformed into corresponding official gene symbols using the Protein Information Resource (PIR, Georgetown University Medical Center, Washington, DC, USA). The PIR is an integrated public bioinformatics resource to support genomic, proteomic, and systems biology research and scientific studies.\(^1\)

**Grouping of PCa proteomics studies**

We divided the selected proteomics literature into two groups: the PCa and benign tissue or cell (P&b) group and the high and low metastatic tendency of PCa (H&L) group. In the P&b group, we compared the protein expression profiles of PCa tissue and BPH, PCa tissue and adjacent benign tissues, and PCa cells and normal prostate cells. The resultant DEPs reveal the changes in related proteins during tumorigenesis. In the H&L PCa group, we compared metastatic PCa (lymph node metastasis, bone metastasis) and localized PCa (cancer limited inside the prostate without metastasis), organ-confined and extracapsular PCa, high Gleason score and low Gleason score, metastatic PCa cell line and cell lines from PCa localized to the prostate, and high and low/poor metastatic PCa cell lines. The resultant DEPs identify changes related to metastasis risk.

**The reported frequency of the DEPs in the literature**

We manually counted the frequency with which each DEP is reported in the literature. If several DEPs were associated with the same official gene symbol in one study, this DEP was considered to have been reported once.

**Functional classification of DEPs**

To examine the biological significance of the DEPs, we carried out a GO enrichment analysis and KEGG pathway analysis of the DEPs from the P&b and H&L groups using the online tool DAVID Bioinformatics Resources 6.7 (The Database for Annotation, Visualization and Integrated Discovery, from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA).\(^1\) To explore the overrepresented biological terms and signaling pathways, we set threshold of false discovery rate to \(P<0.05\).

**Tissue specimens and immunohistochemistry validation**

We selected the DES protein, which was the most commonly reported protein in the P&b group, for further experimental validation. A total of 60 pathological sections were post-operatively collected from each patient at the department of Pathology of the North China University of Science and Technology Affiliated Hospital. These patients were divided into the PCa group, which included 30 patients who had undergone a transurethral resection of the prostate between 2007 and 2014. The mean age, mean serum prostate-specific antigen, and mean prostate volume of this group were 73.1±8.47 years (range: 51–88 years), 50.99±37.02 ng/mL (range: 0.89–102.1 ng/mL), and 58.22±29.13 cm\(^3\) (range: 26.19–142.3 cm\(^3\)), respectively. Postoperative pathologic reports confirmed a diagnosis of prostate adenocarcinoma. None of the patients received radiotherapy or adjuvant therapy before surgery. The patient characteristics are shown in Table 1. The remaining 30 patients were allocated to the BPH group, whose mean age, mean serum prostate-specific antigen, and mean prostate volume were 72.1±8.98 years (range: 55–90 years), 5.38±7.17 ng/mL (range: 0.362–29.86 ng/mL), and 67.68±29.28 cm\(^3\) (range: 23.15–154.77 cm\(^3\)), respectively. All patients underwent transurethral resection of the prostate, and high Gleason score and low Gleason score, metastatic PCa cell line and cell lines from PCa localized to the prostate, and high and low/poor metastatic PCa cell lines. The resultant DEPs identify changes related to metastasis risk.

**Table 1 Demographic and clinical characteristics of patients with prostate cancer**

<table>
<thead>
<tr>
<th>Clinical characteristics</th>
<th>Number of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
</tr>
<tr>
<td>≥70</td>
<td>12</td>
</tr>
<tr>
<td>&lt;70</td>
<td>18</td>
</tr>
<tr>
<td>Prostate volume (cm(^3))</td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>16</td>
</tr>
<tr>
<td>&gt;50</td>
<td>14</td>
</tr>
<tr>
<td>Preoperative PSA (ng/mL)</td>
<td></td>
</tr>
<tr>
<td>≤10</td>
<td>7</td>
</tr>
<tr>
<td>&gt;10</td>
<td>23</td>
</tr>
<tr>
<td>Gleason score</td>
<td></td>
</tr>
<tr>
<td>≤7</td>
<td>13</td>
</tr>
<tr>
<td>8–10</td>
<td>17</td>
</tr>
<tr>
<td>Tumor stage(^a)</td>
<td></td>
</tr>
<tr>
<td>T1–T2</td>
<td>16</td>
</tr>
<tr>
<td>T3–T4</td>
<td>14</td>
</tr>
<tr>
<td>Lymph node metastasis</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>9</td>
</tr>
<tr>
<td>None</td>
<td>21</td>
</tr>
<tr>
<td>Distant metastasis</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>18</td>
</tr>
<tr>
<td>None</td>
<td>12</td>
</tr>
</tbody>
</table>

**Note:** \(^a\)Tumor stage according to American Joint Committee on Cancer (2002).

**Abbreviation:** PSA, prostate-specific antigen.
prostate, and postoperative pathology reports confirmed a diagnosis of BPH. We also excluded individuals with under-active detrusor, unstable bladder, PCa, or a surgery history of BPH. This study was approved by the local ethics committee of the North China University of Science and Technology Affiliated Hospital.

DES antibody was purchased from Abcam (Cambridge, UK), and conjugated antibody PV6000 was purchased from American GBI company (Los Angeles, California, USA). All procedures followed the manufacturer’s instructions. Briefly, 5 μm paraffin sections were deparaffinized in xylene, gradually rehydrated in ethanol, and endogenous enzyme activity was blocked by incubation in 3% hydrogen peroxide for 10 minutes at room temperature. Antigen retrieval was performed using 0.01 mol/L sodium citrate buffer (pH 6.0) at 115°C for 1.5 minutes. The tissue sections were incubated with the primary polyclonal antibody to DES overnight at 4°C. After washing, the sections were incubated with secondary antibody for 25 minutes at 37°C, and the signal was visualized with the Diaminobenzidine map kit. Following counterstaining with hematoxylin, positive immunostaining was scored by two experienced pathologists in a blinded manner. Any discrepant scores were reexamined by both pathologists to reach a consensus score. The immunohistochemistry (IHC) results were scored by examining ten random representative fields at high magnification to observe the positive cells. Each visual field was scored in a semiquantitative manner with the following scale: 1) According to the staining intensity: no coloring =0, light-yellow =1, brown-yellow =2, and chocolate brown =3. 2) According to the staining area: no coloring =0, <25% =1, 25%–50% =2, >50% =3. The staining intensity and staining area were multiplied to produce a weighted score for each case. Cases with scores of 0–3 were designated as “negative”, whereas cases with scores of 4–9 were designated as “positive”.

Statistical analysis

The SPSS 17.0 software package (SPSS Inc., Chicago, IL, USA) was used to conduct the statistical analyses. Quantitative data are expressed as mean ± variance. The χ² test was used to compare the positive rate in the tumor stroma and BPH tissues, and values of P<0.05 were considered statistically significant.

Results

General survey of proteomics data

The data mining and analysis identified a total of 555 DEPs. Of the 555 DEPs, 320 were in the P&b group, 353 were in the H&L PCa group, and 119 appeared in both groups. In the P&b group, 77 DEPs were reported more than twice, and 20 were reported more than three times (Table 2). DES was reported most often (seven times). In the H&L PCa group, 52 proteins were reported more than twice and 13 were reported more than three times (Table 3). Of these proteins, ACPP, HSPA5, and VCL were reported most often (four times). Proteins in the P&b and H&L groups are described in detail in Tables S1 and S2.

GO term enrichment analysis

To gain insight into the biological roles of these DEPs, we performed a GO categories enrichment analysis, which provides a common descriptive framework to functionally annotate and classify gene sets. GO categories are organized into three groups: biological process, cellular component, and molecular function. Using a false discovery rate <0.05 threshold, we identified GO terms for biological processes significantly enriched in response to wounding, regulation of apoptosis, response to organic substance, homeostatic process, cell motion, and cell adhesion in the P&b group in our study. Overrepresented cellular components mainly included extracellular region, intracellular non-membrane-bound organelle, membrane-enclosed lumen, cytoskeleton, vesicle, and mitochondrion. For molecular functions, the enriched GO terms were calcium ion binding, structural molecule activity, cytoskeletal protein binding, identical protein binding, and enzyme inhibitor activity (Figure 1). In the H&L PCa group, we identified GO terms for biological processes significantly enriched in protein localization, protein transport, the regulation of apoptosis, cell adhesion, and homeostasis. Overrepresented cellular components mainly included extracellular region, intracellular non-membrane-bound organelle, cytoskeleton, membrane-enclosed lumen, cell fraction, vesicle, cell projection, and extracellular matrix (ECM). For molecular functions, the enriched GO terms were nucleotide binding, calcium ion binding, cytoskeletal protein binding, structural molecule activity, and identical protein binding (Figure 2).

KEGG pathway analysis of DEPs

We conducted a KEGG pathway analysis for DEPs in P&b and H&L groups. The results show that the over-represented DEPs in the P&b group are mainly related to focal adhesion formation, the complement and coagulation cascades, and the glycolysis/gluconeogenesis signaling pathway (percentage >3%, P-value <0.05) (Table 4).
Bioinformatics analysis of proteomics data in prostate cancer

Table 2 Detailed information about the 20 differentially expressed proteins reported more than three times in the P&b group

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene symbol</th>
<th>Protein name</th>
<th>Frequency of occurrence</th>
<th>Differentially expressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1674</td>
<td>DES</td>
<td>Desmin</td>
<td>7</td>
<td>Downregulation</td>
</tr>
<tr>
<td>308</td>
<td>ANXA5</td>
<td>Annexin A5</td>
<td>4</td>
<td>Upregulation</td>
</tr>
<tr>
<td>2023</td>
<td>ENO1</td>
<td>Enolase 1, (alpha)</td>
<td>4</td>
<td>Upregulation</td>
</tr>
<tr>
<td>3329</td>
<td>HSPD1</td>
<td>Heat shock 60 kDa protein 1</td>
<td>4</td>
<td>Upregulation</td>
</tr>
<tr>
<td>5176</td>
<td>SERPIN1</td>
<td>Serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium-derived factor), member 1</td>
<td>4</td>
<td>1 up, 2 down</td>
</tr>
<tr>
<td>7414</td>
<td>VCL</td>
<td>Vinculin</td>
<td>4</td>
<td>1 up, 2 down</td>
</tr>
<tr>
<td>55</td>
<td>ACPP</td>
<td>Acid phosphatase, prostate</td>
<td>3</td>
<td>Upregulation</td>
</tr>
<tr>
<td>306</td>
<td>ANXA3</td>
<td>Annexin A3</td>
<td>3</td>
<td>Upregulation</td>
</tr>
<tr>
<td>563</td>
<td>AZGP1</td>
<td>Alpha-2-glycoprotein 1, zinc binding</td>
<td>3</td>
<td>Upregulation</td>
</tr>
<tr>
<td>2266</td>
<td>FGG</td>
<td>Fibrinogen gamma chain</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
<tr>
<td>2934</td>
<td>GSN</td>
<td>Gelsolin</td>
<td>3</td>
<td>2 up, 1 down</td>
</tr>
<tr>
<td>2950</td>
<td>GSTP1</td>
<td>Glutathione S-transferase pi 1</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
<tr>
<td>3309</td>
<td>HSPA5</td>
<td>Heat shock 70 kDa protein 5</td>
<td>3</td>
<td>Upregulation</td>
</tr>
<tr>
<td>3856</td>
<td>KRT8</td>
<td>Keratin 8</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
<tr>
<td>5245</td>
<td>PHB</td>
<td>Prohibitin</td>
<td>3</td>
<td>Upregulation</td>
</tr>
<tr>
<td>10631</td>
<td>POSTN</td>
<td>Periostin, osteoblast-specific factor</td>
<td>3</td>
<td>Upregulation</td>
</tr>
<tr>
<td>462</td>
<td>SERPIN1</td>
<td>Serpin peptidase inhibitor, clade C (antithrombin), member 1</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
<tr>
<td>7163</td>
<td>TPD52</td>
<td>Tumor protein DS2</td>
<td>3</td>
<td>Upregulation</td>
</tr>
<tr>
<td>7168</td>
<td>TPM1</td>
<td>Tropomysin 1 (alpha)</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
<tr>
<td>10383</td>
<td>TUBB4B</td>
<td>Tubulin, beta 4B class 1vB</td>
<td>3</td>
<td>Upregulation</td>
</tr>
</tbody>
</table>

Notes: *The listed gene IDs are for the human species. *Information on the protein expression is based on results from the proteomics literature. *This protein was reported once as upregulated and reported twice as downregulated. In this table, the frequency with which proteins are reported is not consistent with the frequency of expression because some studies did not provide expression information.

Abbreviations: P&b, PCa and benign tissues; PCa, prostate cancer; ID, identification.

In the H&L PCa group, the overrepresented DEPs are mainly associated with the regulation of the actin cytoskeleton, focal adhesion formation, ECM–receptor interaction, and the complement and coagulation cascade signaling pathways (percentage >3%, P-value <0.05) (Table 5).

DES protein is downregulated in human PCa tissues

We measured the DES protein levels using IHC in 30 PCa and 30 BPH tissue specimens. Immunohistochemical staining for DES revealed benign tissue cells with concentrated

Table 3 Detailed information about the 13 differentially expressed proteins reported more than three times in the H&L PCa group

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene symbol</th>
<th>Protein name</th>
<th>Frequency of occurrence</th>
<th>Differentially expressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>ACPP</td>
<td>Acid phosphatase, prostate</td>
<td>4</td>
<td>1 up, 2 down</td>
</tr>
<tr>
<td>3309</td>
<td>HSPA5</td>
<td>Heat shock 70 kDa protein 5</td>
<td>4</td>
<td>3 up, 1 down</td>
</tr>
<tr>
<td>7414</td>
<td>VCL</td>
<td>Vinculin</td>
<td>4</td>
<td>2 up, 2 down</td>
</tr>
<tr>
<td>87</td>
<td>ACTN1</td>
<td>Actinin, alpha 1</td>
<td>3</td>
<td>Downregulation</td>
</tr>
<tr>
<td>301</td>
<td>ANXA1</td>
<td>Annexin A1</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
<tr>
<td>563</td>
<td>AZGP1</td>
<td>Alpha-2-glycoprotein 1, zinc binding</td>
<td>3</td>
<td>Upregulation</td>
</tr>
<tr>
<td>1292</td>
<td>COL6A2</td>
<td>Collagen, type VI, alpha 2</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
<tr>
<td>1674</td>
<td>DES</td>
<td>Desmin</td>
<td>3</td>
<td>Downregulation</td>
</tr>
<tr>
<td>2316</td>
<td>FLNA</td>
<td>Filamin A, alpha</td>
<td>3</td>
<td>Downregulation</td>
</tr>
<tr>
<td>4000</td>
<td>LMNA</td>
<td>Lamin A/C</td>
<td>3</td>
<td>2 up, 1 down</td>
</tr>
<tr>
<td>4191</td>
<td>MDH2</td>
<td>Malate dehydrogenase 2, NAD (mitochondrial)</td>
<td>3</td>
<td>2 up, 1 down</td>
</tr>
<tr>
<td>4627</td>
<td>MYH9</td>
<td>Myosin, heavy chain 9, non-muscle</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
<tr>
<td>7094</td>
<td>TNL1</td>
<td>Talin 1</td>
<td>3</td>
<td>1 up, 1 down</td>
</tr>
</tbody>
</table>

Notes: *The listed gene IDs are for the human species. *Information on the protein expression is based on results from the proteomics literature. *This protein was reported once as upregulated and reported twice as downregulated. In this table, the frequency with which proteins are reported is not consistent with the frequency of expression because some studies did not provide expression information.

Abbreviations: PCa, prostate cancer; H&L, high and low PCa metastatic tendencies; ID, identification.
staining in the stroma, and BPH showed strong DES staining; weak staining and no staining were observed in PCa tissue (Figure 3). Specifically, ~73.3% and 10% of BPH and PCa tissues were positive for DES, respectively, and this difference was significant (P < 0.001) (Table 6).

**Discussion**

Proteomics technology plays important roles in the study of PCa occurrence and metastasis by simultaneously monitoring changes in thousands of proteins. However, due to limits in resolution, current proteomics technology can only monitor changes in certain proteins. In addition, it cannot examine many samples due to complex techniques and expensive consumables, which results in unreliable findings. Thus, we conducted general bioinformatics analysis of human PCa occurrence and metastasis using the proteomics literature in PubMed indexed before June 1, 2015. We combined the findings of several proteomics studies. Specifically,
proteomics technologies supplement each other to compensate for shortcomings associated with various methods, which increases the coverage to detect more proteins. In addition, if more than one or similar methods confirm same protein changes, the results are more reliable.

As shown in Tables 2 and 3, we found that some changes in proteins are not consistent. This discrepancy may be due to separation changes or sampling errors caused by differences in modifications, differences in transcripts, or genetic mutations. In the P&b group, DES, ANXA5, ENO1, and other proteins were reported more than three times, and of these DEPs, DES was reported most often (seven times).

Table 4 Results of KEGG pathway analysis of P&b group (all pathways with >3%, P-value < 0.05)

<table>
<thead>
<tr>
<th>KEGG pathway</th>
<th>KEGG entry</th>
<th>Protein count</th>
<th>Percentage</th>
<th>P-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focal adhesion</td>
<td>hsa04510</td>
<td>23</td>
<td>7.5</td>
<td>2.1E-5</td>
</tr>
<tr>
<td>Complement and coagulation</td>
<td>hsa04610</td>
<td>17</td>
<td>5.5</td>
<td>3.6E-8</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>hsa05215</td>
<td>11</td>
<td>3.6</td>
<td>1.6E-2</td>
</tr>
<tr>
<td>Glycolysis/ glutoneogenesis</td>
<td>hsa00010</td>
<td>10</td>
<td>3.2</td>
<td>4.6E-3</td>
</tr>
</tbody>
</table>

Notes: *The percentage of each pathway in the total proteins in P&b group.  
<sup>a</sup>Benjamini method was used to control the false discovery rate (FDR) to correct the P-value.

Abbreviations: P&b, PCA and benign tissues; PCA, prostate cancer; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Figure 2 Results of GO analysis of H&L group (all components with >6%, P-value < 0.05).

Abbreviations: GO, Gene Ontology; PCA, prostate cancer; H&L, high and low PCA metastatic tendencies.
Table 5 Results of KEGG pathway analysis of H&L PCa group (all pathways with ≥3%, P-value <0.05)

<table>
<thead>
<tr>
<th>KEGG pathway</th>
<th>KEGG entry</th>
<th>Proteins count</th>
<th>Percentage</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation of actin cytoskeleton</td>
<td>hsa04810</td>
<td>25</td>
<td>7.4</td>
<td>1.0E-4</td>
</tr>
<tr>
<td>Focal adhesion</td>
<td>hsa04510</td>
<td>24</td>
<td>7.1</td>
<td>5.7E-5</td>
</tr>
<tr>
<td>ECM–receptor interaction</td>
<td>hsa04512</td>
<td>15</td>
<td>4.4</td>
<td>7.5E-5</td>
</tr>
<tr>
<td>Complement and coagulation cascades</td>
<td>hsa04610</td>
<td>11</td>
<td>3.2</td>
<td>3.1E-3</td>
</tr>
</tbody>
</table>

Notes: *The percentage of each pathway in the total proteins in H&L PCa group.

Abbreviations: PCa, prostate cancer; H&L, high and low PCa metastatic tendencies; KEGG, Kyoto Encyclopedia of Genes and Genomes; ECM, extracellular matrix.

This finding suggests that changes in these proteins are closely related to PCa occurrence. These proteins are potential differential diagnosis biomarkers between PCa and benign disease. Among these proteins, SERPINF1 has been shown to be the most potent inhibitor of angiogenesis. In PCa, its expression is downregulated compared with normal prostatic epithelial cells. SERPINF1 can block angiogenesis, induce apoptosis, and prompt the neuroendocrine differentiation of PCa cells. Abnormal methylation of GSTP1 will downregulate the expression of target genes, and promote the progression of PCa. The dynamic interactions between HSPA5 and clusterin (CLU) under endoplasmic reticulum stress conditions may govern CLU trafficking and redistribution to the mitochondria, elucidating how HSPA5 and CLU cooperatively promote survival during treatment stress in PCa. POSTN can induce EMT via the activation of the PI3K/Akt pathway in PCa cells. TP53 promotes cell migration via αvβ3 integrin in PCa cells through activation of the protein kinase B/Akt signaling pathway. Moreover, DES, HSPD1, ANXA3, ANXA5, ENO1, VCL, AZGP1, KRT8, and PHB were reported to be differentially expressed between PCa and normal tissues, whereas the molecular mechanisms of this differential expression need to be further investigated. Currently, the association between FGG, GSN, SERPINC1, TPM1, and TUBB4B and PCa occurrence has not been studied in detail.

In the H&L PCa group, 13 proteins (ACPP, HSPA5, VCL, etc) were reported more than three times. Of these proteins, ACPP, HSPA5, and VCL were the most reported (four times), which suggests that these proteins are potential biomarkers that can be used to differentiate aggressive from non-aggressive PCa. ACPP has already been reported to be expressed in human PCa bone metastases and promote osteoblast differentiation. The upregulated expression of HSPA5 is associated with the development of castration-resistant PCa. VCL overexpression might contribute to PCa progression by enhancing tumor cell proliferation. ANXA1 was shown to regulate EMT in PCa cells. The expression of AZGP1 might be regulated by androgen, and the relationship between AZGP1 and GATA-2 may have functional roles in the transition of PCa cells to a more aggressive phenotype. COL6A2 is likely to be one of the important hyper-methylation genes involved in the pathway from PCa to androgen-independent PCa. FLNA may interact with androgen receptor and may suppress androgen receptor transcriptional activity. LMNA is positively involved in malignant behavior of PCa cells through the PI3K/Akt/PTEN pathway. As a adaptor protein, TLN1 promotes the

Table 6 Comparison of DES expression in benign prostatic hyperplasia tissues and prostate cancer tissues

<table>
<thead>
<tr>
<th>Category</th>
<th>n</th>
<th>Negative</th>
<th>Positive</th>
<th>χ²</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPH</td>
<td>30</td>
<td>8 (26.7%)</td>
<td>22 (73.3%)</td>
<td>24.754</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PCa</td>
<td>30</td>
<td>27 (90%)</td>
<td>3 (10%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: BPH, benign prostatic hyperplasia; PCa, prostate cancer; DES, desmin.

Figure 3 Expression of DES protein in prostate cancer tissues and benign prostatic hyperplasia tissues (>200).
Notes: (A) Positive stromal DES expression in benign prostatic hyperplasia. (B) Negative stromal DES expression in prostate cancer.
Abbreviation: DES, desmin.
activation of integrins, which further promote tumor cell motility by forming focal adhesions. To the best of our knowledge, the association between MDH2 or MYH9 and PCAs has not yet been investigated, which warrants further investigation. In addition, ACPP, HSPA5, VCL, AZGP1, and DES were also reported more than three times in the P&b group, suggesting that these proteins may be involved in PCa occurrence and metastasis.

A protein cancer biomarker is a protein measured in body fluids or in tissues that could reflect the presence of cancer and indicate its aggressiveness, staging, and response to therapy. Protein cancer biomarkers can be divided into several categories: diagnostic screening biomarkers, prognostic biomarkers, and stratification biomarkers. Proteomics, together with the innovative high-throughput technologies, might be a highly promising way to identify new biomarkers for both detection and tailoring therapy.

This study is to extract some DEPs that had been repeatedly reported from previous PCa-related proteomics literature. To further validate the hypothesis that frequently reported proteins can be used as potential differential diagnosis biomarkers, we examined the expression of DES protein, which was the most commonly reported protein in the P&b group, using IHC. The results showed that compared with BPH tissue, DES was significantly downregulated in cancer tissues. This expression is consistent with the expression levels of this protein reported in the proteomics literature in Table 2, indicating that DES can be used as a potential biomarker to detect PCa and differentiate benign disease. Furthermore, this result is consistent with the experimental results obtained by Wu et al from clinical samples. Moreover, Ayala et al also proposed that DES can be used as a late-stage PCa smooth muscle marker. Specifically, its expression is also downregulated during PCa progression; patients exhibiting lower DES expression levels had a shorter disease-free period. This finding agrees with the high frequency with which DES was reported in the H&L PCa group, suggesting that DES can also be used as a potential biomarker to differentiate aggressive PCa from non-aggressive PCa. DES is a muscle-specific type III intermediate filament essential for proper muscular structure and function.

Now several studies have shown that DES is a highly sensitive marker for endothelial cell differentiation and tumor invasiveness in several types of cancers, including colon cancer, gastrointestinal stromal tumors, endometrial carcinoma, and embryonal sarcomas. Moreover, some researches asserted that DES keeps continuous high expression in the formation process of early tumor capillaries. Therefore, DES was further characterized in this study to confirm the potential diagnostic and prognostic value for PCa. The aforementioned DEPs all repeatedly appeared in the proteomics literature, which suggests that these proteins are potential biomarkers of PCa. Therefore, these DEPs warrant further study to determine the true meaning of the differential expression.

A GO analysis of the DEPs in both the P&b and H&L groups revealed that proteins related to the regulation of apoptosis, cell adhesion, cell motion, homeostasis, and the response to wounding appeared in both groups. This finding suggests that these biological processes may be closely associated with PCa. The abnormal regulation of apoptosis, loss of cell–cell adhesion, cell motility via the EMT of epithelial cells, and collective cell motility can promote PCa occurrence and metastasis. However, few studies have examined the roles of homeostasis and response to wounding in the molecular biology of PCa. The homeostatic balance between tumor suppressor proteins and oncoproteins in normal prostate epithelia is apparently altered in cancer cells. Changes in Ca homeostasis can also trigger the development of PCa. In addition, some data support that the stromal cells behave nearly identically in response to a wound or a tumor. Accordingly, the role of stromal cells in tissue homeostasis, particularly in wound healing, is functionally linked to the role of stromal cells in the PCa-generated stromal response. Therefore, further studies are required to take advantage of the stromal response in formulating new therapeutic approaches. In the H&L PCa group, enrichment in processes related to protein localization and protein transport suggested that these processes may play important roles in PCa development and metastasis. We searched the literature for DEPs that had been reported more than three times in the P&b and H&L groups to study the molecular mechanisms by which these proteins are involved in PCa. We found that many of the biological processes involved overlap with the results of the GO analysis. To illustrate their possible combined influence on PCa as a whole, these proteins were mapped in the background of all the involved biological processes in Figure 4. However, this approach provides far from a comprehensive snapshot of all the PCa-related proteins because of the limited number of proteins identified in this study.

By studying the DEPs in the P&b and H&L groups via the KEGG pathway analysis, we found that focal adhesion formation and the complement and coagulation cascades pathways simultaneously appear in both groups (Tables 4 and 5). Focal adhesions are special regions at which cells
attach to the ECM, and previous studies have indicated that FAK is frequently overexpressed and overactive in PCa. Through the activation of major oncogenic pathways, FAK promotes the growth, survival, migration, metastasis, and androgen-independence of prostate tumors.

In recent years, some studies have reported that the complement and coagulation cascades are associated with PCa. Hong et al. reported that complement C1q may induce PCa apoptosis by activating WOX1 and destabilizing cell adhesion. The down-regulation of C1q enhances tumorigenesis due to WOX1 inactivation. Therefore, complement C1q is an important factor in the occurrence of PCa that warrants further study.

In addition, the glycolysis/gluconeogenesis pathway was enriched in the P&b group. PPARγ, CAV1, and LRP6 have been reported to be able to affect glycolysis in PCa cells to slow cell growth and cause apoptosis by interrupting energy metabolism. Table 5 shows that the regulation of the actin cytoskeleton and ECM–receptor interaction pathway are enriched in the H&L PCa group. The enhanced motility of cancer cells due to the remodeling of the actin cytoskeleton is crucial in the process of cancer cell invasion and metastasis. Multiple studies have also demonstrated changes in expression of ECM molecules in advanced PCa tumor samples, such as increased expression of bone sialoprotein and CDH-11 or decreased collagen type VII expression. However, few studies have examined PCa and ECM–receptor interaction, and this interaction warrants further examination.

---

**Figure 4** The map of the differentially expressed proteins reported more than three times in the context of all the involved biological processes.

**Notes:** Italic indicates downregulation. The rest is upregulation. Underline: proteins from P&b group; double wavy lines: proteins from H&L group; bold underline: proteins from both groups. ‘—’ indicates inhibition or degradation; ‘—’ indicates transition or flow.

**Abbreviations:** CAMs, cell adhesion molecules; EMT, epithelial–mesenchymal transition; ER, endoplasmic reticulum; P&b, PCa and benign tissues; PCa, prostate cancer; TGF, transforming growth factor; H&L, high and low PCa metastatic tendencies.
Conclusion
In conclusion, we conducted a bioinformatics analysis of the proteomics literature on PCa and identified several PCa-associated proteins. The repeatedly reported DEPs between the P&b and H&L groups may serve as potential biomarkers to discriminate PCa from benign tissue or identify metastasis risk. The experimental detection of one of these proteins in clinically relevant sets of samples verified our hypothesis. However, some repeatedly reported proteins lacked detailed studies. The overrepresentation of biological processes and signaling pathways increased our understanding of the molecular mechanisms of PCa carcinogenesis and metastasis and may even provide additional targets for clinical diagnosis and treatment.

Acknowledgments
This study was supported by the National Natural Science Foundation of China (Grant No 81100518), Youth Foundation of Higher Education Scientific Research of Hebei Province (Grant No QN2014013), the Social Science Fund Project of Hebei Province (Grant No HB12SH030), and the Tangshan Research and Development Program (Grant No 111302112b).

Disclosure
The authors report no conflicts of interest in this work.

References


