Pathway crosstalk analysis based on protein-protein network analysis in prostate cancer

J.-M. WANG\textsuperscript{1,2}, J.-T. WU\textsuperscript{2}, D.-K. SUN\textsuperscript{2}, P. ZHANG\textsuperscript{2}, L. WANG\textsuperscript{2}

\textsuperscript{1}Department of Urology, Qilu Hospital, Shandong University, Jinan, China
\textsuperscript{2}Department of Urology, Yantai Yuhuangding Hospital, Yantai, Shandong Province, China

Abstract. - BACKGROUND: Prostate cancer (PCa) is a highly prevalent disease in older men of the western world and overall greatly affects the quality of life of elderly people.
AIM: Understanding the mechanism of prostate cancer onset and metastasis is the key to treating this disease successfully and increasing survivability.
MATERIALS AND METHODS: In this study, we constructed crosstalk networks among prostate cancer (primary and metastatic) related pathways by integrating protein-protein interactions (PPI) and KEGG (Kyoto encyclopedia of genes and genomes) pathways information. Total 11 pathways crosstalk with each other in primary prostate cancer and 7 pathways crosstalk with each other in metastatic prostate cancer.
RESULTS: Among these pathways, Notch signaling pathway and chemokine signaling pathway were found regulate multiple processes during prostate cancer progression.
CONCLUSIONS: Results from these studies will provide the groundwork for a combination therapy approach targeting multiple pathways which will likely be more effective than targeting one pathway alone.

Key Words: Prostate cancer, Protein-protein network, Pathway crosstalk.

Introduction

Prostate cancer (PCa) is a highly prevalent disease in older men of the western world\textsuperscript{1,2}. It is estimated that 241,740 men will be diagnosed with and 28,170 men will die of cancer of the prostate in 2012 in the United States\textsuperscript{3}. Although the age-adjusted rate of cancer deaths has decreased steadily in the past 10 years, PCa remains the second leading cause of cancer deaths in men after lung cancer\textsuperscript{4}. The morbidity and mortality of prostate cancer is principal caused of its propensity to metastasize to other tissue, such as lung, liver and bone\textsuperscript{5,6}. Hence, understanding the mechanism of prostate cancer onset and metastasis is the key to treating this disease successfully and increasing survivability\textsuperscript{7}.

Currently, there are several popular hypothesis of prostate cancer pathogenesis\textsuperscript{8-10}. However, the genetic mechanisms of ovarian cancer are far from being clear. Some genes have been found to be aberrant in a significant proportion of prostate cancer, including glutathione S-transferase P1 (GSTP1), PTEN, TP53, and androgen receptor (AR)\textsuperscript{11}. The loss function of PTEN (phosphatase and tensin homolog) results in up-regulation of the Akt/mTOR (protein kinase B/mammalian target of rapamycin) signaling pathway in prostate cancer, primarily through activation of Akt1 (protein kinase B)\textsuperscript{12-15}. MAPK (mitogen-activated-protein-kinase) signaling is also frequently activated in prostate cancer, particularly in advanced disease, and is often coordinately deregulated together with Akt signaling\textsuperscript{12,16-18}. Wnt and TGF-β (transforming growth factor-β) signaling pathways are also of potential importance for prostate cancer pathogenesis, based on the presence of β-catenin mutations\textsuperscript{19}. However, these investigations rarely consider the potential relationships among these pathways. The disease complexity is coming from not only the cooperations of proteins in the form of pathways but also the interactions of pathways, i.e., the crosstalks of these \textsuperscript{20,21}.

Recently, systems biology approaches such as network-based methods have been successfully applied to elucidate the mechanism of diseases\textsuperscript{22,23}. Crosstalk between pathways provides more complex nonlinear responses to combinations of dysfunctions\textsuperscript{24}. From the systematic perspective, analysis of ovarian cancer related biomolecular interaction networks will improve the understanding of the complexity of molecular pathways underlying ovarian cancer and will help to uncover the dynamic processes of disease progression.

Corresponding Author: Wang Lin, MD; e-mail: wanglin2688@hotmail.com
In this study, we sought to explore the molecular mechanism of prostate cancer onset and metastasis using a network-based analysis by integrating protein-protein interactions and gene expression profiles. The availability and integration of high-throughput gene expression data and the genome-wide protein-protein interaction may shed new lights on prostate cancer study.

Materials and Methods

Affymetrix Microarray Data

We extracted the gene expression profile data on prostate cancer patients with normal controls from\(^{25}\), which were deposited in NCBI GEO (National Center for Biotechnology Information Gene Expression Omnibus) (http://www.ncbi.nlm.nih.gov/geo/) database (ID: GSE3325) based on the Affymetrix Human Genome U133 plus 2.0 Array. Total 19 chips were available for further analysis, including 13 individual benign prostate, primary and metastatic prostate cancer samples and 6 pooled samples from benign, primary or metastatic prostate cancer tissues\(^{25}\).

Protein-Protein Interaction (PPI) data

The Human Protein Reference Database (HPRD)\(^{26}\) is a protein database accessible through the internet. The Biological General Repository for Interaction Datasets (BioGRID)\(^{27}\) is a curated biological database of protein-protein and genetic interactions. Total 326119 unique PPI pairs were collected in which 39240 pairs are from HPRD and 379426 pairs are from BioGRID. We constructed an ensemble protein-protein interaction network by integrating the PPI data collected from the two above PPI databases in human.

Differential Expressed Genes (DEGs) Analysis

We preprocessed the CEL source files by RMA (robust multarray averaging) algorithm with defaulted parameters in R bioconductor package\(^{28}\). The RMA method is a mathematical technique used to obtain variance stabilization and reduce discrepancies in hybridization patterns that might result from variables in target amplification, hybridization conditions, staining, or probe array lots\(^{29}\). Probe sets were mapped to NCBI entrez genes using DAVID (Database for Annotation, Visualization and Integrated Discovery)\(^{30}\). If there are multiple probe sets that correspond to the same gene, the expression values of those probe sets are averaged. The limma method\(^{31}\) was used to identify DEGs. The DEGs only with the fold change value larger than 2 or less than -2 were selected.

Significance Analysis of Pathway Under PPI Data

To determine the co-expressed significance of a gene pair in disease cases, we used the Pearson correlated coefficient test to calculate the \(p\)-value.

Then we mapped those \(p\)-values to the nodes and edges in the PPI network. The following formula is used to define a function as the combination of statistical significance of an interaction by a scoring scheme. The detailed description could be seen in Liu et al\(^{32}\).

\[
S(e) = f[\text{diff}(x), \text{corr}(x, y), (\text{diff}(y))]
\]

\[
= -2 \sum_{i=1}^{k} \log_e (p_i)
\]

The \(\text{diff}(x)\) and \(\text{diff}(y)\) are differential expression assessments of gene \(x\) and gene \(y\), respectively. \(\text{corr}(x,y)\) represents their correlation between gene \(x\) and gene \(y\). \(f\) is a general data integration method that can handle multiple data sources differing in statistical power. Where \(k = 3\), \(p1\) and \(p2\) are the \(p\)-values of differential expression of two nodes, \(p3\) is the \(p\)-value of their co-expression.

To define the significance of a pathway \(P, Sp\), we summarize all the scores of edges \(S(e)\) of every pathway.

\[
Sp = \sum_{e \in P} S(e)
\]

To estimate a \(p\)-value for significance of this pathway, we iteratively compute similar scores 10\(^5\) times on randomly generated pathways of the same size as that of pathway \(P\). The frequency of scores that are larger than \(Sp\) is used as the significant \(p\)-value of pathway \(P\). We considered the pathway with the \(p\)-value < 0.05 as the significant pathway.

Pathway Crosstalk Analysis Based on the GO (Gene Ontology) Enrichment Analysis

Analysis of crosstalk of relationships among pathways is then investigated, especially that with overlapping GO-ID of two significant pathway analysis results.

The functional enrichment among proteins in one pathway is defined as:

\[
Pathway crosstalk analysis based on protein-protein network analysis in prostate cancer

\[ P = 1 - \sum_{i=0}^{k-1} \left( \binom{i}{f} \binom{n-f}{m-i} \right) \binom{n}{m} \]

where \( n \) is the number of nodes in the network, \( f \) is the number of proteins annotated with a particular GO function, \( m \) is the number of proteins involved in the pathway and \( k \) is the frequency of the GO-ID. We identified the GO function enrichment of the pathways respectively with the \( p \)-value < 0.05.

Results

Differentially Expressed Genes Selection

To get differentially expressed genes between primary prostate cancer and metastatic prostate cancer, we obtained publicly available microarray dataset GSE3325 from GEO. The fold change value larger than 2 or less than -2 were chosen as cutoff criterion. After microarray data analysis, total 5847 genes were identified as DEGs between benign prostate tissue and primary prostate cancer, and total 2026 genes were identified as DEGs between benign prostate tissue and metastatic prostate cancer. There were a total of 977 overlapping DEGs.

Significant Pathways in Primary Prostate Cancer and Metastatic Prostate Cancer

Pathway can provide an alternative way to relax the significance threshold applied to single genes and may lead to a better biological interpretation\(^{21}\). To identify the relevant pathways changed in primary prostate cancer and metastatic prostate cancer, we used a statistical approach on pathway level. The significance analysis of pathway is based on the protein-protein interaction database. The impact analysis method yielded many significant pathways containing Notch signaling pathway, regulation of autophagy, chemokine signaling pathway and so on (Tables I and II).

Crosstalk of GO Relationships Among Pathways

We considered the pathway crosstalk between significant pathways detected by the overlapping GO-IDs. For detailed analysis of the crosstalk between pathways, we used the hypergeometric test to find the significant GO-ID in each pathway with the \( p \)-value less than 0.05, respectively. We found that 11 pathways crosstalk with each other in primary prostate cancer (Figure 2) and 7 pathways crosstalk with each other in metastatic prostate cancer (Figure 3). The results of the significant GO-ID in part of the pathways are used to construct the connection among pathways. From the significant GO enrichments, we know

<table>
<thead>
<tr>
<th>Path_ID</th>
<th>Size</th>
<th>Node</th>
<th>Edge</th>
<th>( p )-value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa04330</td>
<td>48</td>
<td>39</td>
<td>795</td>
<td>0</td>
<td>Notch signaling pathway</td>
</tr>
<tr>
<td>hsa05322</td>
<td>143</td>
<td>24</td>
<td>158</td>
<td>0</td>
<td>Systemic lupus erythematosus</td>
</tr>
<tr>
<td>hsa04140</td>
<td>36</td>
<td>14</td>
<td>90</td>
<td>0</td>
<td>Regulation of autophagy</td>
</tr>
<tr>
<td>hsa04062</td>
<td>190</td>
<td>127</td>
<td>834</td>
<td>2.00E-05</td>
<td>Chemokine signaling pathway</td>
</tr>
<tr>
<td>hsa04060</td>
<td>266</td>
<td>113</td>
<td>464</td>
<td>2.00E-05</td>
<td>Cytokine-cytokine receptor interaction</td>
</tr>
<tr>
<td>hsa04012</td>
<td>88</td>
<td>11</td>
<td>25</td>
<td>3.00E-05</td>
<td>ErbB signaling pathway</td>
</tr>
<tr>
<td>hsa03420</td>
<td>45</td>
<td>35</td>
<td>240</td>
<td>9.00E-05</td>
<td>Nucleotide excision repair</td>
</tr>
<tr>
<td>hsa00534</td>
<td>27</td>
<td>4</td>
<td>14</td>
<td>0.00139</td>
<td>Glycosaminoglycan biosynthesis – heparan sulfate</td>
</tr>
<tr>
<td>hsa04114</td>
<td>115</td>
<td>82</td>
<td>1486</td>
<td>0.00342</td>
<td>Oocyte meiosis</td>
</tr>
<tr>
<td>hsa04310</td>
<td>152</td>
<td>52</td>
<td>133</td>
<td>0.00589</td>
<td>Wnt signaling pathway</td>
</tr>
<tr>
<td>hsa05218</td>
<td>72</td>
<td>34</td>
<td>219</td>
<td>0.00604</td>
<td>Melanoma</td>
</tr>
<tr>
<td>hsa00460</td>
<td>8</td>
<td>2</td>
<td>6</td>
<td>0.01813</td>
<td>Cytosine acid metabolism</td>
</tr>
<tr>
<td>hsa04920</td>
<td>68</td>
<td>10</td>
<td>32</td>
<td>0.0271</td>
<td>Adipocytokine signaling pathway</td>
</tr>
<tr>
<td>hsa04623</td>
<td>57</td>
<td>5</td>
<td>11</td>
<td>0.03587</td>
<td>Cytosolic DNA-sensing pathway</td>
</tr>
<tr>
<td>hsa04621</td>
<td>63</td>
<td>43</td>
<td>318</td>
<td>0.03706</td>
<td>NOD-like receptor signaling pathway</td>
</tr>
<tr>
<td>hsa03030</td>
<td>37</td>
<td>2</td>
<td>7</td>
<td>0.04462</td>
<td>DNA replication</td>
</tr>
</tbody>
</table>

“Size” is the number of genes contained in KEGG gene sets. “Edge” and “Node” represent the number of edges and nodes of these pathways in our protein-protein interaction network with gene expression information. “\( p \)-value” gives the \( p \)-value of dysregulation score of every pathway in the cluster. “Description” gives the pathway name.
the crosstalk of GO biological processes during the transcription and immune response among the pathways. In Figure 2, hsa04330 (Notch signaling pathway), hsa04062 (chemokine signaling pathway), hsa04310 (Wnt signaling pathway) are hub nodes in the network, which suggest that dysfunctional of these pathways play an important role in the progression from benign prostate tissue to primary prostate cancer. In Figure 3, hsa04062 (chemokine signaling pathway) and hsa04330 (Notch signaling pathway) are hub nodes in the network, which suggest that dysfunctional of these pathways play an important role in the progression from benign prostate tissue to metastatic prostate cancer.

Some of the identified crosstalks between the pathways are consistent with our knowledge for prostate cancer; some of them provide valuable alternatives for the mechanism of prostate cancer, especially from the pathway relationship perspective. The interactions of these pathways provide more insights for the prostate cancer progression. Besides, we also identified the biological processes enrichments underlying these interacted pathways. The GO functional linkages of these pathways provided more implications for their dysfunctional crosstalk.

From the result of crosstalk network between significant pathways in primary prostate cancer, we found that several immunity and cancer related pathways were identified crosstalk with each other in our method, such as nucleotide oligomerization domain (NOD)-like receptor signaling pathway, Notch signaling pathway, ErbB (erythroblastic leukemia viral oncogene homolog) signaling pathway, Wnt signaling pathway and so on. Especially, hsa04330 (Notch signaling pathway), hsa04062 (chemokine signaling pathway), hsa04310 (Wnt signaling pathway) are hub nodes in the network, which suggest that dysfunction of these pathways play an important role in the progression from benign prostate tissue to primary prostate cancer.

The Notch signaling pathway is known to be crucial in communication between adjacent cells, which involves important processes during embryonic and adult life. Notch signals are required for normal prostate development and homeostasis, and abnormalities in Notch signaling may be critical during the development of prostate cancer. The Notch pathway proposed as potential therapeutic target for several types of cancer, including prostate cancer. Chemokines are members of a superfamily of chemotactic cy-
Chemokines, initially recognized as critical mediators of the inflammatory response by regulating recruitment of cells from both the innate and adaptive immune systems to the site of infection. Induction of chemotaxis has been observed in a number of carcinoma cell lines including those of breast and prostate. Chemokines were also shown to regulate growth of certain carcinoma cell lines. During the transitions from normal to PCa, a number of chemokines and chemokine receptors display variations in their expression, such as CXCL8, CXCL12, and CCL2. Wnt signaling pathway often involves in Ca²⁺ signaling, which leads to transient increases in cytoplasmic free calcium that subsequently activates the calcium kinase and the phosphatase calcineurin. Wnt pathway dysfunction is an important component of prostate tumorigenesis and offers a variety of treatment targets.

Hsa04062 (chemokine signaling pathway) and hsa04330 (Notch signaling pathway) are hub nodes in the metastatic prostate cancer crosstalk network, which suggest that dysfunctional of these pathways also play an important role in the progression from benign prostate tissue to metastatic prostate cancer. The role of chemokine signaling pathway and Notch signaling pathway...
in the prostate cancer metastases have been widely documented. Emerging studies show that chemokines are critical for cancer progression and indicate complex and diverse functions in the tumor microenvironment. Chemokines and their receptors show differential expression along with PCa progression. Compare with less aggressive cells, the more aggressive types of PCa cell lines express higher levels of CCR2. Several studies showed that Jagged1, a Notch receptor ligand, is significantly more highly expressed in metastatic prostate cancer, suggesting the biological significance of Notch signaling in prostate cancer progression.

We also identified the biological processes enrichments underlying these interacted pathways. The results of the top significant GO terms in part of the pathways are shown under the network. In primary prostate cancer, the three significant pathways are linked by the biological processes transcription, DNA-dependent (GO: 0006351) and proteolysis (GO: 0006508). The functions correspond to the main process in prostate tumorigenesis. In metastatic prostate cancer, the significant functions of pathways provide a flow of transporting of proteins (GO: 0015031). From the significant GO enrichments, we know the crosstalk of GO biological processes during the prostate cancer development between the significant pathways.

**Conclusions**

We constructed two crosstalk networks among primary prostate cancer related pathways and metastatic prostate cancer related pathways. We also identified the biological processes enrichments underlying these interacted pathways. Several signaling pathways which play crucial roles in prostate cancer were identified crosstalk with each other in our method. Total 11 pathways crosstalk with each other in primary prostate cancer and 7 pathways crosstalk with each other in metastatic prostate cancer. Among these pathways, Notch signaling pathway and chemokine signaling pathway are hub nodes in both primary prostate cancer and metastatic prostate cancer. They regulate multiple processes during prostate cancer progression. Therefore, inactivation of

---

**Figure 3.** Crosstalk network between significant pathways in metastatic prostate cancer. The blue lines indicate that the two pathways are connected with the same GO-ID. The green nodes stand for the significant pathways.
Notch signaling pathway and chemokine signaling pathway by innovative strategies could be potential targeted approach for the treatment of prostate cancer. Results from these studies will provide the groundwork for a combination therapy approach targeting multiple pathways which will likely be more effective than targeting one pathway alone.

References


41) VADAY GG, PEEHL DM, KADAM PA, LAWRENCE DM. Expression of CCL5 (RANTES) and CCR5 in prostate cancer. Prostate 2006; 66: 124-134.


