Identification of significant pathways in gastric cancer based on protein-protein interaction networks and cluster analysis

Kongwang Hu\textsuperscript{1} and Feihu Chen\textsuperscript{2}

\textsuperscript{1}Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Anhui, P.R. China.
\textsuperscript{2}School of Pharmacology, Anhui Medical University, Anhui, P.R. China.

Abstract

Gastric cancer is one of the most common and lethal cancers worldwide. However, despite its clinical importance, the regulatory mechanisms involved in the aggressiveness of this cancer are still poorly understood. A better understanding of the biology, genetics and molecular mechanisms of gastric cancer would be useful in developing novel targeted approaches for treating this disease. In this study we used protein-protein interaction networks and cluster analysis to comprehensively investigate the cellular pathways involved in gastric cancer. A primary immunodeficiency pathway, focal adhesion, ECM-receptor interactions and the metabolism of xenobiotics by cytochrome P450 were identified as four important pathways associated with the progression of gastric cancer. The genes in these pathways, e.g., ZAP70, IGLL1, CD79A, COL6A3, COL3A1, COL1A1, CYP2C18 and CYP2C9, may be considered as potential therapeutic targets for gastric cancer.

Key words: graph clustering, pathway crosstalk, protein-protein interaction network.

Received: March 12, 2012; Accepted: May 4, 2012.

Introduction

Gastric cancer is one of the most common malignancies worldwide (Lin \textit{et al.}, 2007b). Surgical resection is the only effective treatment for this cancer, although current surgical therapeutic strategies are far from optimal and most patients are diagnosed with late-stage disease when surgical intervention is of limited use (D’Ugo \textit{et al.}, 2009). Chemotherapy has been applied as a neoadjuvant treatment to improve the curative resection rate or to achieve long-term survival in patients with unresectable gastric cancer. The prognosis, however, is still unsatisfactory, with an overall five-year survival rate of 24% (Kanai \textit{et al.}, 2003). Hence, there is an urgent need for new therapeutic strategies.

Recently, several molecular alterations involving various pathways have been implicated in the development and late-stage progression/metastasis of gastric cancer. For example, there is emerging evidence that the Wnt signaling pathway may contribute to gastric carcinogenesis by stimulating the migration and invasion of gastric cancer cells (Kurayoshi \textit{et al.}, 2006). Persons with germ-line mutations in the APC tumor suppressor gene have a 10-fold increased risk of developing gastric cancer when compared with normal persons (Offerhaus \textit{et al.}, 1992). β-catenin is frequently mutated in gastric cancer (Clements \textit{et al.}, 2002). In addition, frizzled receptor E3 (FzE3) is over-expressed in 75% of gastric carcinoma tissues and secreted frizzled related protein (hsFRP) is down-regulated in 16%, suggesting that alterations in FzE3 and hsFRP expression are frequent in this pathology (To \textit{et al.}, 2001). Activation of the hedgehog pathway is another important mechanism associated with aggressive gastric cancer. The sonic hedgehog (Shh) transcript is restricted to cancer tissue whereas Gli1 and human patched gene 1 (PTCH1) are expressed in cancer cells and the surrounding stroma. The treatment of gastric cancer cells with 3-keto-N-aminoethylamino-caproyldihydrocinnamoyl-cyclopamine, a hedgehog signaling inhibitor, decreases the expression of Gli1 and PTCH1 and results in cell growth inhibition and apoptosis (Ma \textit{et al.}, 2005). The high recombinant Shh-induced migration and invasiveness of gastric cancer cells is mediated by tissue growth factor-beta (TGF-β) acting through the ALK5-Smad3 pathway (Yoo \textit{et al.}, 2008). The expression of lysyl oxidase-like 2 (LOXL2), which can promote tumor cell invasion via the Src kinase/focal adhesion kinase (Src/FAK) pathway, is markedly increased in gastric cancer (Peng \textit{et al.}, 2009). The loss of embryonic liver fodrin (ELF) can disrupt TGF-β-mediated signaling by interfering with the localization of Smad3 and Smad4 and leads to the development of gastric cancer (Kim \textit{et al.}, 2006).
An increased concentration of BMP-2 strongly enhances the motility and invasiveness in gastric cancer cells. The stimulation of BMP-2 in gastric cancer cells induces a full epithelial-mesenchymal transition (EMT) characterized by Snail induction, E-cadherin reorganization and the down-regulation and up-regulation of mesenchymal and invasiveness markers through the activation of phosphatidylinositol 3 (PI-3) kinase/Akt (Kang et al., 2010). Cysteine-rich 61 (Cyr61) may contribute to the progression of gastric cancer by promoting tumor cell motility/invasion through the up-regulation of cyclooxygenase-2 (COX-2) in an integrin αvβ3/NF-kB-dependent manner. Interleukin-6 induces gastric cancer cell line AGS cell invasion through activation of the c-Src/RhoA/ROCK signaling pathway (Lin et al., 2007a).

The use of high-throughput approaches to dissect the molecular mechanisms and pathways that regulate the progress of gastric cancer is still comparatively rare. In this study, we used microarray data, protein-protein interaction (PPI) networks and cluster graph analysis to identify significant pathways involved in the development of gastric cancer. The characterization of genes and pathways involved in gastric cancer should be useful in identifying potential targets for the development of novel strategies for treating gastric carcinoma.

Data and Methods

Data sources

The KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa, 2002) datasets were downloaded on February 19, 2011, at which time they contained 211 pathways and 5,385 genes. The PPI data were collected from the HPRD (Human Protein Reference Database) (Keshava Prasad et al., 2007), MINT (Molecular INTeraction Database) (Chatr-aryamontri et al., 2007) and BIOGRID (Biological General Repository for Interaction Datasets) (Stark et al., 2011). A total of 21,978 unique PPI pairs were obtained, of which 21,353 were from HPRD, 8,830 were from MINT and 19,243 were from BIOGRID. An ensemble PPI network was constructed by integrating three of the above PPI databases for humans, with at least two PPI databases being used to form an intersection (the PPI data are provided as Table S1 in Supplementary Material).

The gene expression profile data were accessed at the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) data repository using the accession number GSE2685. Samples of gastric cancer tissue and corresponding adjacent noncancerous tissue were obtained with the informed consent of patients who underwent gastrectomy at Jichi Medical College Hospital (Tochigi, Japan) (Hippo et al., 2002). Twenty-two gastric cancer tissue samples and eight noncancerous gastric tissue samples were analyzed with oligonucleotide microarrays (GeneChip Hu-GeneFL array; Affymetrix, Santa Clara, CA).

Analysis of significant pathways based on cluster graph analysis

The Limma eBayes analysis (Smyth, 2004) was used to assess the differential expression status of each gene. Background intensities were adjusted and the original expression datasets from all conditions were processed into expression estimates using the robust multiarray average (RMA) method (D’Souza et al., 2008) with the default settings implemented in R (version 2.12.1) (Gentleman et al., 2004); this was followed by construction of the linear model. The empirical Bayes approach was used to further justify these estimators; this process is equivalent to shrinking the estimated sample variances towards a pooled estimate and yields a far more stable inference when the number of arrays is small (Smyth, 2004). At least a two-fold change in expression and a p value of < 0.05 were considered as the threshold for defining differentially expressed genes (DEGs). Spearman’s rank correlation (r) was used to assess the association between different DEGs. The level of significance was set at r > 0.75 and the false discovery rate (FDR) at < 0.05 (Strimmer, 2008). All statistical tests were done using the R program.

To identify co-expressed groups we used DPClus (Altaf-Ul-Amin et al., 2006). DPClus is a cluster graph algorithm that can extract densely connected nodes as a cluster and is based on density tracking and peripheral tracking of clusters. In this study, we used the overlapping-mode of the DPClus settings. For this analysis, we used the parameters “cluster property” (cp), a density value of 0.5 and a minimum cluster size of 5 (Fukushima et al., 2011). DAVID software (Huang da et al., 2008) was used for pathway enrichment analysis with p < 0.05 selected as the threshold for gene clusters based on their hypergeometric distribution.

Analysis of significant pathways and pathway crosstalk based on PPI networks

Pathway crosstalk was defined as those pathways that had overlapping genes and edges. “Overlapping genes” meant that both of the pathways included these genes whereas “overlapping edges” meant that both pathways included the PPI interaction edges. Liu et al. (2010) have provided a detailed analysis of crosstalk relationships. The significance of a co-expressed gene pair in gastric cancer was assessed using Pearson’s correlation coefficient and the corresponding p values, with the latter being mapped to the nodes and edges in the PPI network. The final identification of significant pathways was based on the extent of overlap of the pathways identified by the two methods (cluster graph analysis and PPI networks).
Results

Identification of significant pathways based on screening for differentially expressed genes and cluster graph analysis

A publicly available microarray dataset (GSE2685) was downloaded from GEO and screened for DEGs. In the microarray analysis, 723 genes with a fold change > 2 and p < 0.05 were identified as DEGs using the limma eBayes method. Based on the cutoffs established for r (> 0.75) and FDR (< 0.05) a correlation network was constructed that included 1032 relationships among 364 DEGs. At r > 0.75, DPClus identified 22 clusters that ranged in size from 5 to 24 genes, with each cluster being connected to neighboring clusters (Figure 1). The significance of the clusters was assessed by examining the over-represented pathways in these clusters (also known as pathway enrichment analysis). Table 1 shows the results of this analysis based the cluster graphs in Figure 1. Only clusters 1, 2, 3, 4, 8 and 19 contained enriched pathways.

Primary immunodeficiency (hsa05340) enriched in cluster 1 was connected with the metabolism of xenobiotics by cytochrome P450 (hsa00980), linoleic acid metabolism (hsa00591) and retinol metabolism (hsa00830), which were enriched in cluster 4 (Figure 1). Primary immunodeficiency (hsa05340) was also connected with the B cell receptor signaling pathway (hsa04662) that was enriched in cluster 19. Cluster 2, in which ECM-receptor interaction (hsa04512) and focal adhesion (hsa04510) were enriched, was indirectly connected with cluster 3 that included cell cycle (hsa04110), oocyte meiosis (hsa04114) and DNA replication (hsa03030).

Identification of significant pathways and pathway crosstalk based on PPI networks

Twenty significant pathways with p < 0.05 were detected using the KEGG pathways and PPI datasets (Table 2). Further analysis of these pathways revealed only 11 cases of crosstalk that involved nine significant pathways (Figure 2). Primary immunodeficiency (hsa05340) showed crosstalk with the ribosome (hsa03010) and chemokine signaling pathway (hsa04062). More importantly, cluster graph analysis and PPI networks identified primary immunodeficiency (hsa05340), focal adhesion (hsa04510), metabolism of xenobiotics by cytochrome P450 (hsa00980) and ECM-receptor interaction (hsa04512) as overlapping significant pathways with

---

Table 1 - Clusters showing pathway enrichment.

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Description</th>
<th>Count</th>
<th>p-value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>hsa05340</td>
<td>Primary immunodeficiency</td>
<td>3</td>
<td>0.0029</td>
<td>0.0675</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>hsa04512</td>
<td>ECM-receptor interaction</td>
<td>4</td>
<td>0.0003</td>
<td>0.0034</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>hsa04510</td>
<td>Focal adhesion</td>
<td>4</td>
<td>0.0043</td>
<td>0.0212</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>hsa04110</td>
<td>Cell cycle</td>
<td>5</td>
<td>0.0002</td>
<td>0.0003</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>hsa04114</td>
<td>Oocyte meiosis</td>
<td>3</td>
<td>0.0119</td>
<td>0.0805</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>hsa03030</td>
<td>DNA replication</td>
<td>2</td>
<td>0.0553</td>
<td>0.2331</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>hsa00591</td>
<td>Linoleic acid metabolism</td>
<td>2</td>
<td>0.0432</td>
<td>0.6968</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>hsa00830</td>
<td>Retinol metabolism</td>
<td>2</td>
<td>0.0819</td>
<td>0.6846</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>hsa00590</td>
<td>Arachidonic acid metabolism</td>
<td>2</td>
<td>0.0848</td>
<td>0.5497</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>hsa00980</td>
<td>Metabolism of xenobiotics by cytochrome P450</td>
<td>2</td>
<td>0.0906</td>
<td>0.4734</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>hsa00982</td>
<td>Drug metabolism</td>
<td>2</td>
<td>0.0935</td>
<td>0.4116</td>
</tr>
<tr>
<td>Cluster 8</td>
<td>hsa00190</td>
<td>Oxidative phosphorylation</td>
<td>2</td>
<td>0.0984</td>
<td>0.6801</td>
</tr>
<tr>
<td>Cluster 19</td>
<td>hsa04662</td>
<td>B cell receptor signaling pathway</td>
<td>2</td>
<td>0.0293</td>
<td>0.1878</td>
</tr>
</tbody>
</table>

Term represents the pathway identification (ID), Description is the pathway symbol and Count is the number of enriched pathways. The p value is the probability of obtaining a test statistic. The smaller the p value, the greater the number of enriched pathways. The False discovery rate (FDR) is a statistical method used to correct for multiple comparisons in multiple hypotheses testing; the smaller the FDR, the greater the correctness. ECM – extracellular matrix.
considerable crosstalk, e.g., between focal adhesion (hsa04510) and ECM-receptor interaction (hsa04512) based on the PPI network. Cluster graph analysis showed that both of these two pathways were enriched in cluster 2.

The table below lists the pathways showing significant crosstalk.

<table>
<thead>
<tr>
<th>Pathway ID</th>
<th>Description</th>
<th>Size</th>
<th>Node</th>
<th>Edge</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa00071</td>
<td>Fatty acid metabolism</td>
<td>42</td>
<td>2</td>
<td>2</td>
<td>0.0038</td>
</tr>
<tr>
<td>hsa00280</td>
<td>Valine, leucine and isoleucine degradation</td>
<td>44</td>
<td>2</td>
<td>2</td>
<td>0.0064</td>
</tr>
<tr>
<td>hsa00520</td>
<td>Amino sugar and nucleotide sugar metabolism</td>
<td>45</td>
<td>2</td>
<td>2</td>
<td>0.0107</td>
</tr>
<tr>
<td>hsa00534</td>
<td>Glycosaminoglycan biosynthesis – heparan sulfate</td>
<td>26</td>
<td>2</td>
<td>3</td>
<td>0.0301</td>
</tr>
<tr>
<td>hsa00910</td>
<td>Nitrogen metabolism</td>
<td>23</td>
<td>2</td>
<td>4</td>
<td>0.0113</td>
</tr>
<tr>
<td>hsa00980</td>
<td>Metabolism of xenobiotics by cytochrome P450</td>
<td>70</td>
<td>4</td>
<td>8</td>
<td>0.0133</td>
</tr>
<tr>
<td>hsa03010</td>
<td>Ribosome</td>
<td>88</td>
<td>19</td>
<td>39</td>
<td>0.0438</td>
</tr>
<tr>
<td>hsa03060</td>
<td>Protein export</td>
<td>24</td>
<td>2</td>
<td>2</td>
<td>0.0113</td>
</tr>
<tr>
<td>hsa03420</td>
<td>Nucleotide excision repair</td>
<td>44</td>
<td>25</td>
<td>58</td>
<td>0</td>
</tr>
<tr>
<td>hsa04012</td>
<td>ErbB signaling pathway</td>
<td>87</td>
<td>3</td>
<td>3</td>
<td>0.0120</td>
</tr>
<tr>
<td>hsa04062</td>
<td>Chemokine signaling pathway</td>
<td>189</td>
<td>61</td>
<td>176</td>
<td>0</td>
</tr>
<tr>
<td>hsa04310</td>
<td>Wnt signaling pathway</td>
<td>151</td>
<td>5</td>
<td>6</td>
<td>0.0453</td>
</tr>
<tr>
<td>hsa04510</td>
<td>Focal adhesion</td>
<td>201</td>
<td>3</td>
<td>2</td>
<td>0.0038</td>
</tr>
<tr>
<td>hsa04740</td>
<td>Olfactory transduction</td>
<td>389</td>
<td>2</td>
<td>2</td>
<td>0.0384</td>
</tr>
<tr>
<td>hsa04930</td>
<td>Type II diabetes mellitus</td>
<td>47</td>
<td>5</td>
<td>6</td>
<td>0.0181</td>
</tr>
<tr>
<td>hsa04964</td>
<td>Proximal tubule bicarbonate reclamation</td>
<td>23</td>
<td>3</td>
<td>5</td>
<td>0.0145</td>
</tr>
<tr>
<td>hsa04512</td>
<td>ECM-receptor interaction</td>
<td>132</td>
<td>7</td>
<td>23</td>
<td>0.0307</td>
</tr>
<tr>
<td>hsa05144</td>
<td>Malaria</td>
<td>51</td>
<td>5</td>
<td>27</td>
<td>0.0480</td>
</tr>
<tr>
<td>hsa05322</td>
<td>Systemic lupus erythematosus</td>
<td>142</td>
<td>9</td>
<td>16</td>
<td>0.0180</td>
</tr>
<tr>
<td>hsa05340</td>
<td>Primary immunodeficiency</td>
<td>35</td>
<td>3</td>
<td>3</td>
<td>0.0332</td>
</tr>
</tbody>
</table>

*Description* refers to the pathway name. *Size* is the number of genes contained in the KEGG gene sets. *Edge* and *Node* represent the number of edges and nodes of the pathways in the protein-protein interaction network that contain gene expression information. The *p-value* indicates the dysregulation score for each pathway. ECM – extracellular matrix.

Discussion

Graph clustering or PPI-based pathway analysis (Hwang et al., 2008) has been successfully used to identify the underlying mechanisms associated with diseases. In this study, we used the same strategy to identify DEGs associated with gastric cancer and predict their underlying molecular mechanisms. Our cluster analysis showed that primary immunodeficiency (hsa05340) enriched in cluster 1 interacted not only with the metabolism of xenobiotics by cytochrome P450 (hsa00980), linoleic acid metabolism (hsa00591) and retinol metabolism (hsa00830) that were enriched in cluster 4, but also with the B cell receptor signaling pathway (hsa04662) enriched in cluster 19. These results indicated that the primary immunodeficiency pathway has several important roles in gastric cancer. Although additional pathways were observed in PPI-based pathway analysis, the primary immunodeficiency pathway still retained important roles in gastric cancer through crosstalk with the ribosome (hsa03010) and chemokine signaling pathway (hsa04062). In addition, focal adhesion (hsa04510), metabolism of xenobiotics by cytochrome P450 (hsa00980) and ECM-receptor interaction (hsa04512) were all enriched in the two methods used to assess pathway enrichment. Based on these findings, we suggest that the genes of the primary immunodeficiency...
pathway, focal adhesion, ECM-receptor interaction and the metabolism of xenobiotics by cytochrome P450 (hsa00980) are potentially important therapeutic targets for gastric cancer. These pathways are discussed below in greater detail.

Primary immunodeficiencies are a heterogeneous group of disorders that affect cellular and humoral immunity or non-specific host defense mechanisms mediated by complement proteins and by cells such as phagocytes and natural killer cells. These immune system disorders cause increased susceptibility to malignancy. For example, patients with common variable immunodeficiency, the second most prevalent primary immunodeficiency in adults, have a 10-fold increased risk of gastric cancer (Dhalla et al., 2011). Patients in advanced stages of gastric cancer frequently suffer from cell-mediated immunodeficiency, such as the inhibition of interleukin-2 production, the main cytokine that modulates the cell-mediated immune response, and a decrease in the total and T lymphocyte counts (Romano et al., 2003). In addition, the absolute number of T-regulatory lymphocytes (Tregs; CD4+CD25+Foxp3+) is significantly lower in gastric cancer patients than in normal individuals (Szczepanik et al., 2011).

Disorders in the regulation of humoral immunity also have a significant effect on the development of gastric cancer. For example, a significant increase in IgG Fc fucosylation has been observed in stages II and III of gastric cancer (Kodar et al., 2012). The widespread expression of CD40, a member of the tumor necrosis factor receptor superfamily, reflects the central role of CD40 in regulating humoral immunity and host defense. The stimulation of CD40 in gastric carcinoma makes cells less vulnerable to apoptosis induced by Fas or chemotherapy and increases cell motility (Yamaguchi et al., 2003).

ZAP70 (zeta-chain (TCR) associated protein kinase 70 kDa), IGLL1 (immunoglobulin lambda-like polypeptide 1) and CD79A (CD79a molecule, immunoglobulin-associated alpha) were enriched in the primary immunodeficiency pathway. ZAP70 may be involved in T-cell-mediated immunodeficiency. ZAP-70 ectopic expression leads to enhanced B cell receptor signaling after IgM stimulation and increased expression of CCR7 (chemokine [C-C motif] receptor 7), predominantly via ERK1/2, thereby enhancing the response to and migration towards CCL21 (chemokine [C-C motif] ligand 21). In addition, cellular subsets with high ZAP-70 expression in chronic lymphocytic leukemia show increased expression of adhesion molecules and chemokine receptors (Calpe et al., 2011). IGLL1 and CD79A are associated with B cell-mediated immunodeficiency. Mutations in IGLL1 and CD79A can result in B cell deficiency and few or no γ-globulins or antibodies are produced (Storlazzi et al., 2002; Wang et al., 2002). These alterations may promote the metastasis of gastric cancer since an anti-Wnt5a antibody suppresses the Wnt5a-dependent internalization of receptors. This in turn prevents the metastasis of gastric cancer cells by inhibiting the activation of Rac1 (ras-related C3 botulinum toxin substrate 1 [rho family, small GTP binding protein Rac1]) and the expression of laminin γ2 (Hanaki et al., 2012). Based on these findings, we conclude that the primary immunodeficiency pathway may affect the progress of gastric cancer by inhibiting T lymphocyte proliferation and antibody production by B lymphocytes, or by enhancing the expression of adhesion molecules and chemokine receptors.

The interaction between tumor cells and extracellular matrix (ECM) components such as laminin, fibronectin and collagen, has a crucial role in tumor invasion and metastasis. This interaction is facilitated by adhesion receptors such as integrins. Consequently, ECM-receptor interactions and the focal adhesion pathway may be involved in cancer metastasis. Collagen is the major constituent of the tumor ECM and several types of collagens have been implicated in the focal adhesion and ECM-receptor interaction pathways in gastric carcinoma (Yin et al., 2009). Watanabe et al. (1995) reported greater deposition of type III collagen at the periphery of poorly differentiated gastric cancer tissue compared with more central locations.

Microarray studies have shown the enhanced expression of several collagen genes (COL1A1, 1A2, 3A1, 4A1, 4A2, 4A6, 5A2, 6A3, 7A1, 9A3, 11A1 and 18A1) in the endothelium of gastric cancer tissue compared with normal endothelium (Hippo et al., 2002; Oue et al., 2004). The most up-regulated genes in gastric cancer, such as COL1A1, 1A2, 3A1, 4A1 and 4A2, are associated with cell adhesion or migration and the ECM. COL4A6, 6A3, 17A1 and 18A1 are also associated with cell adhesion, COL1A1 with cell growth and/or maintenance, and COL1A2 and 6A3 with the `ECM-receptor interaction' pathway (Yasui et al., 2004). In agreement with previous studies, COL6A3, 3A1 and 1A1 were found to be involved in focal adhesion and ECM-receptor interaction pathways. These findings suggest that targeting these genes with RNA interference could decrease the collagen content of the ECM in gastric carcinoma and reduce cell proliferation and migration. This diversity of collagens suggests that each type is associated with some aspect of gastric cancer. The identification of collagens as potential therapeutic targets will require a more complete understanding of their expression and interactions in gastric cancer.

Cytochromes P450 (CYP) are a multi-gene family of constitutive and inducible heme-containing enzymes with a crucial role in the metabolism of xenobiotics, including many potential carcinogens and various anti-cancer drugs. CYP P450s have a central role in chemical carcinogenesis and are involved in tumor initiation and promotion because they can activate or deactivate most carcinogens. Furthermore, CYP P450s can influence the response of established tumors to anti-cancer drugs by metabolizing these drugs in tumor cells (Ding and Kaminsky, 2003). The expression of major isoforms of P450, such as CYP1A and CYP3A, is en-
hanced in gastric cancer, with CYP1A being enhanced in 51% of cases and CYP3A in 28% (Murray et al., 1998). CYP2C9, CYP3A7 and CYP3A5 that participate in drug metabolism are down-regulated in gastric cancer. In Helicobacter pylori-positive Japanese, poor metabolizers via CYP2C19 have an increased risk of developing gastric cancer, especially the diffuse type (Sugimoto et al., 2005). In Chinese with gastric cancer the frequency of poor metabolizers via CYP2C19 is 31.8% (Shi and Chen, 2004).

As shown here, CYP2C18 and CYP2C9 were associated with the development of gastric cancer through the metabolism of xenobiotics by cytochrome P450, arachidonic acid metabolism, retinol metabolism and the linoleic acid metabolism pathway. CYP2C9 is one of the predominant epoxidegenase isoforms involved in the metabolism of arachidonic acid into 12-epoxyeicosatetraenoic acid (EEF). CYP2C9 epoxidegenases are upregulated in human tumors and promote tumor progression and metastasis (Xu et al., 2011). Retinol may influence gastric carcinogenesis through its essential role in controlling cell proliferation and differentiation. High intakes of retinol from foods or a combination of foods and supplements are associated with a lower risk of gastric cancer (Larsson et al., 2007). CYP2C18 and CYP2C9 are related to retinol metabolism in human through their ability to transform retinol into 4-OH-retinoic acid and 18-OH-retinoic acid (Marill et al., 2000). These all-trans-retinoic acids are associated with G0/G1 phase arrest and decreased VEGF expression in human gastric cancer cell lines (Zhang et al., 2007). However, dietary linoleic acid stimulates the invasion and peritoneal metastasis of gastric carcinoma cells through COX-catalyzed metabolism and the activation of ERK (Matsuoka et al., 2010). CYP2C9 is involved in linoleic acid epoxidation and the major product of this reaction is leukotxin that increases oxidative stress and subsequent pro-inflammatory events (Viswanathan et al., 2003), leading to tumor cell progression. We therefore suggest that P450 family genes are involved in gastric cancer by metabolizing exogenous anti-cancer drugs, stimulating arachidonic acid and linoleic acid metabolism and inhibiting retinol metabolism.

In conclusion, the results described here show that changes in the primary immunodeficiency pathway, focal adhesion, ECM-receptor interactions and the metabolism of xenobiotics by cytochrome P450 may be associated with gastric cancer. A number of candidate genes (ZAP70, IGLL1, CD79A, COL6A3, COL3A1, COL1A1, CYP2C18 and CYP2C9) that may be involved in gastric cancer were also identified. Overall, these findings shed new light on the biology of gastric cancer and indicate new avenues for future research.

Acknowledgments

This work was supported by the Anhui Provincial Natural Science Funding for Key Projects in 2010 (grant no. KJ2010A171) and the Medicine Research Projects Schedule of the Province Health Bureau in 2009 (grant no. 09C157).

References


Internet Resources

R program: http://www.r-project.org/ (November 11, 2011).


Supplementary Material
The following online material is available for this article:
Table S1 - PPI data.
This material is available as part of the online article from http://www.scielo.br/gmb.

Associate Editor: Carlos F.M. Menck

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.