Identification of molecular characteristics induced by radiotherapy in rectal cancer based on microarray data

CHANG GE¹, MENGXIA WU¹, GUIFANG CHEN¹, GUANYING YU², DEHUI JI¹ and SHAOZHAO WANG¹

Departments of ¹Anorectal Surgery and ²Gastrointestinal Surgery, Jinan Central Hospital Affiliated to Shandong University, Jinan, Shandong 250013, P.R. China

Received July 27, 2015; Accepted December 6, 2016

DOI: 10.3892/ol.2017.5750

Abstract. The present study aimed to reveal the molecular characteristics induced by radiotherapy in rectal cancer at the transcriptome level. Microarray data (ID, GSE26027) downloaded from the Gene Expression Omnibus database were re-analyzed to identify differentially expressed genes (DEGs) between rectal cancer tissues during and prior to radiotherapy. The DEGs were then inputted into the database for annotation, visualization and integrated discovery, an online tool to perform enrichment analyses, and into the search tool for the retrieval of interacting genes/proteins database to identify protein-protein interactions (PPIs). Subsequently, a PPI network was constructed, which was screened for densely connected modules. Furthermore, protein domain enrichment analysis was performed. In total, 690 DEGs, including 179 upregulated and 511 downregulated DEGs, were found in rectal cancer tissues during and prior to radiotherapy. The upregulated DEGs were significantly enriched in 'positive regulation of transport' and 'regulation of cardiac muscle contraction', while the downregulated DEGs were most markedly enriched in 'cell migration', 'cell-cell signaling', 'extracellular matrix organization' and 'blood vessel development', including prostaglandin-endoperoxide synthase 2, transforming growth factor β-induced, 68 kDa endothelin receptor type A, brain-derived neurotrophic factor, TIMP metalloproteinase inhibitor 1, and serpin family E member 1, which were the top 6 hub nodes in the PPI network. Furthermore, 2 protein domains were significantly enriched by PPI modules, including: The collagen triple helix repeat (CTHR) family members collagen type (COL) 5A2, COL9A3, COL6A3, COL21A1, COL5A3, COL11A1, COL7A1 and CTHR-containing-1; and the olfactory receptor family (OR) members OR7E24, OR7A17, OR6A2, OR1F1, OR10H3 and OR7A10. A total of 7 upregulated DEGs were characterized as tumor suppressor genes, and 8 downregulated DEGs were characterized as oncogenes. The biological processes or protein domains enriched by upregulated or downregulated DEGs may improve the understanding of molecular characteristics in response to radiotherapy.

Introduction

In 2014 ~40,000 new cases of rectal cancer were diagnosed in the USA, with an incidence of 12.3/100,000 ranking as the second leading cause of cancer-associated mortality (1). Up to 40% patients with rectal cancer develop metastatic rectal cancer (2). The management of rectal cancer has developed into an integrated approach involving surgery, chemotherapy, radiotherapy and biological therapy (3), which has improved the 5-year survival rate of patients with rectal cancer (4).

Radiotherapy is commonly used as an adjuvant therapy for treating rectal cancer (5). Preoperative radiotherapy has been suggested as an essential aspect of the treatment options for rectal cancer, and can effectively reduce the local recurrence rate of rectal cancer (6). A number of molecular markers, including p53, p27, p21, epidermal growth factor receptor, B-cell lymphoma 2 (Bcl-2)/Bcl-2-associated X protein and human phosphatidylethanolamine-binding protein 4 have been revealed to participate in the rectal cancer response to radiotherapy (7-9). However, despite the potential of radiotherapy in terms of treating rectal cancer, the treatment may cause severe side effects such as bowel dysfunction and obstruction, incontinence, sexual dysfunction, unspecified infection, abdominal pain, nausea and secondary cancer (6,10,11). Thus, a comprehensive evaluation of radiotherapy for the treatment of rectal cancer is required in order to develop effective therapeutic approaches for patients with rectal cancer. Previously, transcriptomic changes in response to radiotherapy in patients with rectal cancer using microarray data were revealed (12). However, only differential expression analysis and function analysis were performed in the aforementioned study, and the microarray data can be investigated reveal the molecular response to radiotherapy. The present study re-analyzed the gene expression profiles (12) of rectal cancer during and prior to radiotherapy to screen out differentially expressed genes (DEGs) with thresholds of the absolute value of log₂ fold change (FC) >1 and P<0.05 which were different compared

Correspondence to: Dr Chang Ge, Department of Anorectal Surgery, Jinan Central Hospital Affiliated to Shandong University, 105 Jiefang Road, Lixia, Jinan, Shandong 250013, P.R. China
E-mail: gechang@126.com

Key words: rectal cancer, radiotherapy, differentially expressed genes, enrichment analyses, protein-protein interaction, module
with the criteria of FC (particularly the ratio ‘gene expression during radiotherapy/gene expression prior to radiotherapy’) >2.5 or <0.4 and false discovery rate <0.11. In addition, enrichment analyses and functional annotation were subsequently performed for DEGs to reveal their biological relevance. Furthermore, protein-protein interactions (PPIs) were analyzed and visualized into a PPI network to identify hub nodes, and the modules were mined from the PPI network.

Materials and methods

**Microarray data.** The gene expression profiles (ID, GSE26027) (12) were obtained from the Gene Expression Omnibus database (National Institutes of Health, Bethesda, MD, USA), which were previously produced based on the platform of Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix, Inc., Santa Clara, CA, USA). The biopsies for the microarray analysis were obtained from 6 patients with locally advanced rectal cancer prior to radiotherapy and 1 h subsequent to a dose of 7.2 Gy by the 4th fraction (1.8 Gy/fraction) during radiotherapy. All patients enrolled in the experiment provided written informed consent approved by the University of Nantes Institutional Review Board for human studies (Nantes, France) (12).

**Identification of DEGs.** Raw expression data were pre-processed by conducting the following analyses: Background correction; quantile normalization; probe summarization; and generation of the gene expression matrix. In order to take a global view on transcriptome changes, DEGs between biopsies taken during and prior to radiotherapy were screened out using thresholds of the absolute value of log₂ FC>1 and P<0.05 using the t test implemented in Linear Models for Microarray Analysis, (Version 3.0, http://www.bioconductor.org/packages/release/bioc/html/limma.html) package (13,14).

**Enrichment analyses.** To identify biological functions involving DEGs, the identified DEGs were submitted to the Database for Annotation, Visualization and Integrated Discovery (DAVID), version 6.7 (15), to identify Gene Ontology (http://geneontology.org/) functions (16,17) and Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg/) pathways (18). P<0.01 was set as the cut-off criterion.

**Functional annotation.** Transcription factors were identified among DEGs using the TRANSFAC database (version 10.4, http://www.gene-regulation.com/pub/databases.html), which records eukaryotic transcription factors (TFs), their binding sites and DNA binding profiles (19). Oncogenes and
tumor suppressor genes (TSGs) were determined according to the TSGene database (20) and Tumor Associated Gene database (21,22).

**PPI network construction and module mining.** As the interactions between proteins are important for biological metabolism, the DEGs were inputted into the Search Tool for the Retrieval of Interacting Genes (STRING) version 9.1 (23). The interaction pairs were screened from the experimental data, textual mining results, co-expression analysis findings and other interacting records in the STRING database. The interaction pairs were then visualized by constructing a PPI network via Cytoscape version 3.0 (24). Hub nodes, specifically, hub DEGs, were subsequently identified in the PPI network.

Furthermore, densely connected modules were identified from the PPI network, using the threshold of P<0.001 and the plugin Clustering with Overlapping Neighborhood Expansion (25) in Cytoscape software.

**Protein domain enrichment analysis.** The InterPro database (EMBL-EBI, Hinxton, UK) provides family and domain information for the functional analysis of protein sequences (26). Based on the InterPro database, protein domain enrichment analysis was conducted in the present study for the densely connected modules using DAVID (15), classifying proteins into families and predicting their domains and important sites.

### Results

#### DEGs

With the cut-offs of the absolute value of log₂ FC >1 and P<0.05, the present study identified 690 DEGs in rectal cancer tissues during radiotherapy compared to those prior to radiotherapy, including 179 upregulated and 511 downregulated DEGs. There were a greater number of downregulated DEGs than upregulated DEGs, indicating a tendency to downregulate in response to radiotherapy.

#### Results of enrichment analyses

By inputting the DEGs into the online tool DAVID, the upregulated DEGs were significantly enriched in ‘positive regulation of transport’ (P=2.06E-05) and ‘regulation of cardiac muscle contraction’ (P=8.27E-05) (Table I). However, the downregulated DEGs were most significantly enriched in ‘cell migration’ (P=1.44E-09): Prostaglandin-endoperoxide synthase 2 (PTGS2), endothelin receptor type A (EDNRA), TIMP metallopeptidase inhibitor 1 (TIMP 1) and serpin family E member 1 (SERPINE1), ‘cell-cell signaling’: PTGS2 and brain-derived neurotrophic factor (BDNF), ‘extracellular matrix organization’: Transforming growth factor β-induced, 68 kDa (TGFBI), TIMP1 and SERPINE1, and ‘blood vessel development’: PTGS2, TGFBI, EDNRA, LSAMP, NDRG4, LPL, S100A2, IL24.

#### Results of functional annotation

In total, 7 upregulated DEGs such as bone morphogenetic protein 10 (BMP10) and 8 downregulated DEGs such as twist family BHLH transcription factor 1 (TWIST1) were, respectively, identified as tumor suppressor genes and oncogenes (Table II). In addition, 6 upregulated DEGs and 20 downregulated DEGs such as TWIST1 were revealed to be TFs (Table II). However, 24 downregulated DEGs were characterized as TSGs, and 1 upregulated gene, Moloney sarcoma oncogene, was identified as an oncogene (Table II).

#### PPI network and modules

The interaction pairs of DEGs were inputted into Cytoscape and a PPI network was obtained containing 271 nodes and 458 interaction pairs (Fig. 1). The top 6 hub DEGs with a high degree score were identified, including PTGS2 (degree=19), TGFBI (degree=19), EDNRA (degree=14), BDNF (degree=13), TIMP1 (degree=13) and SERPINE1 (degree=13).

Furthermore, using the threshold P<0.01, a total of 3 modules were mined from the PPI network, including modules 1 (P=7.42E-05), 2 (P=1.37E-04) and 3 (P=6.31E-04);
Fig. 2). DEGs with a high degree score in modules were found, including ENDR3 in module 1 and collagen type V α2 (COL5A2) in module 2.

**Results of protein domain enrichment analysis.** By performing protein domain enrichment analysis, module 2 was revealed to be most significantly enriched in the protein domain ‘collagen triple helix repeat’ (P=7.60E-15), e.g. COL5A2 and collagen triple helix repeat containing 1 (CTHRC1), while module 3 was markedly associated with protein domain ‘olfactory receptor’ (P=1.09E-8; Table III). No protein domain was enriched by the proteins in module 1 (Table III).

**Figure 1.** Protein-protein interaction network of DEGs. Red and green nodes indicate upregulated and downregulated DEGs, respectively, while lines indicate interactions between DEGs. The size of each node indicates the number of interactions with other DEGs. DEGs, differentially expressed genes.

**Figure 2.** Densely connected modules screened from the protein-protein interaction network. Red and green nodes indicate upregulated and downregulated DEGs, respectively, while lines indicate interactions between DEGs. The size of each node indicates the number of interactions with other DEGs. M1, module 1; M2, module 2; M3, module 3; DEGs, differentially expressed genes.
Table III. Results of protein domain enrichment analysis on the differentially expressed genes.

<table>
<thead>
<tr>
<th>Term</th>
<th>P-value</th>
<th>Count</th>
<th>Gene list</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1</td>
<td>-</td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>Module 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPR008160: Collagen triple helix repeat</td>
<td>7.60E-15</td>
<td>8</td>
<td>COL5A2, COL9A3, COL6A3, COL21A1, COL5A3, COL11A1, COL7A1, CTHRC1</td>
</tr>
<tr>
<td>IPR000885: Fibrillar collagen, C-terminal 1.78E-05</td>
<td>3</td>
<td>COL5A2, COL5A3, COL11A1</td>
<td></td>
</tr>
<tr>
<td>IPR003129: Laminin G, thrombospondin-type, N-terminal 7.44E-05</td>
<td>3</td>
<td>COL21A1, COL5A3, COL11A1</td>
<td></td>
</tr>
<tr>
<td>IPR002035: von Willebrand factor, type A 9.75E-04</td>
<td>3</td>
<td>COL7A1, COL21A1, COL6A3</td>
<td></td>
</tr>
<tr>
<td>Module 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPR000725: Olfactory receptor</td>
<td>1.09E-08</td>
<td>6</td>
<td>OR7E24, OR7A17, OR6A2, OR1F1, OR10H3, OR7A10</td>
</tr>
<tr>
<td>IPR017452: GPCR, rhodopsin-like superfamily 1.55E-07</td>
<td>6</td>
<td>OR7E24, OR7A17, OR6A2, OR1F1, OR10H3, OR7A10</td>
<td></td>
</tr>
<tr>
<td>IPR000276: 7 transmembrane GPCR, rhodopsin-like 1.56E-07</td>
<td>6</td>
<td>OR7E24, OR7A17, OR6A2, OR1F1, OR10H3, OR7A10</td>
<td></td>
</tr>
</tbody>
</table>

GPCR, G-protein coupled receptor; COL, collagen type; OR, olfactory receptor.

Discussion

Currently, radiotherapy is an important treatment approach for rectal cancer, although the treatment may induce body damage (10,27). The present study provides a comprehensive view on transcriptome changes in response to radiotherapy by re-analyzing previous microarray data (12). As a result, 179 upregulated and 511 downregulated DEGs were identified with the criteria of the absolute value of log2 FC >1 and P<0.05. In a previous study, Supiot et al (12) found that 31 and 6 DEGs were upregulated and downregulated, respectively, by radiotherapy with the criteria of FC (specifically, the ratio ‘gene expression during radiotherapy/gene expression prior to radiotherapy’) >2.5 or <0.4 and false discovery rate <0.11. The differences of DEG numbers between the present and previous study may be due to the distinct P-value thresholds. The criteria used in the present study are commonly utilized in differential expression analysis, therefore the DEGs identified are reasonable.

Subsequent to the enrichment analysis, the upregulated DEGs were mainly associated with the regulation of transport and cardiac muscle contraction. Ion transport is important for cardiac muscle contraction (28), which can be affected by cranial radiotherapy in individuals that survived childhood cancer (29). In the present study, the DEGs involved in the aforementioned processes were significantly upregulated by radiotherapy, suggesting the effects of radiotherapy on cardiac muscle contraction. In addition, the top 4 biological processes significantly enriched by downregulated DEGs were cell migration, cell-cell signaling, extracellular matrix organization and blood vessel development. It has been reported that cell migration and cell-cell communication are important for rectal cancer progression and metastasis (30,31), extracellular matrix organization and angiogenesis (32). The downregulated hub DEGs in the PPI network, including PTGS2, TGFBI, EDNRA, BDNF, TIMP1 and SERPINE1, were revealed to be involved in the aforementioned processes. The level of PTGS2 expression, which exhibits predictive usage for managing rectal cancer (33), is increased by preoperative radiotherapy and involved in local relapse (34). TGFBI serves as a linker protein, and the overexpression of the protein contributes to colorectal cancer development (35). The upregulation of BDNF, SERPINE1 and TIMP1 are associated with colorectal cancer metastasis (36-38). Therefore, the present study hypothesizes that the downregulation of the aforementioned DEGs may alter the biological processes associated with rectal cancer progression and metastasis, which were important molecular responses to radiotherapy in rectal cancer. However, the downregulation of EDNRA, a tumor suppressor, contributes to colorectal cancer progression (21). A decrease in the expression level of BDNF is associated with irradiation-induced hippocampal neurogenesis impairment in Sprague Dawley rats (39). Therefore, radiotherapy in rectal cancer may also cause side effects via regulating the aforementioned genes.

In addition, 3 densely connected modules were identified in the PPI network, and the proteins in 2 of the modules were, respectively, enriched in the protein domains ‘collagen triple helix repeat’, e.g. CTHRC1 and COL5A2, and ‘olfactory receptor’. Collagen is a major component of the interstitial extracellular matrix, which performs a role in cellular proliferation, differentiation, apoptosis, migration and carcinogenesis (40). Due to the interaction with the collagen triple helix repeat domain, CTHRC1 promotes rectal cancer invasion and metastasis with vascular endothelial growth factor C (41), while COL5A2 is co-expressed with COL11A1 in colorectal carcinomas and associated with malignancy in colorectal cancer (40). Numerous types of collagen were also significantly enriched in this domain, and thus may be involved in extracellular matrix change in response to radiotherapy (42,43). Olfactory receptors in the olfactory epithelium are usually overexpressed in tumors and promote cellular.
invasion and metastasis (44). In the present study, the olfactory receptor-associated DEGs in module 3 were downregulated by radiotherapy and thus may contribute towards preventing rectal cancer progression.

In the present study, 7 TSGs such as BMP10 were upregulated and 8 oncogenes including TWIST1 were downregulated in rectal cancer tissues subsequent to radiotherapy. BMP10 is a tumor suppressor in urothelial, prostate and breast cancer, and the overexpression of the gene inhibits the growth, adhesion, migration and invasion of cancer cells (45,46). As an oncogene involved in the epithelial-to-mesenchymal transition, TWIST1 possesses angiogenic, invasive, oncogenic and drug-resistant properties in human tumors. In colorectal cancer, TWIST1 is overexpressed and correlates with lymph node metastasis, overall survival and disease-free survival rates (47). Therefore, radiotherapy may benefit patients with rectal cancer via upregulating TSGs and downregulating oncogenes. However, although 24 TSGs were downregulated, 1 oncogene was upregulated, indicating that radiotherapy in rectal cancer may also promote cancer development.

In conclusion, the present study revealed various molecular responses to radiotherapy in rectal cancer. Cardiac muscle contraction may be affected by radiotherapy in patients with rectal cancer, while metastasis-associated biological processes may be prevented by radiotherapy. In addition, enriched protein domains such as the collagen triple helix repeat and olfactory receptors suggested that radiotherapy degrades the extracellular matrix and promotes metastasis in rectal cancer. In comparison with a previous study (12) that used the dataset GSE26027, the present study identified more DEGs using reasonable criteria, found a potential hub of DEGs by constructing a PPI network and identified TFs, oncogenes and TSGs in DEGs, providing information on the effects of radiotherapy at a molecular level. However, the hypotheses of the present study were formed based on the microarray data of only 6 patients. Therefore, additional exploration and validation is required, and the present authors are planning to use more samples to verify the DEGs with a high degree score in PPI network in future studies, and the annotated TFs, oncogenes and TSGs.

References

10. Birgisson H, Páhlman L, Gunnarsson U and Glimelius B: Reduced expression of the gene inhibits the growth, adhesion, migration and invasion of cancer cells (45,46). As an oncogene involved in the epithelial-to-mesenchymal transition, TWIST1 possesses angiogenic, invasive, oncogenic and drug-resistant properties in human tumors. In colorectal cancer, TWIST1 is overexpressed and correlates with lymph node metastasis, overall survival and disease-free survival rates (47). Therefore, radiotherapy may benefit patients with rectal cancer via upregulating TSGs and downregulating oncogenes. However, although 24 TSGs were downregulated, 1 oncogene was upregulated, indicating that radiotherapy in rectal cancer may also promote cancer development.

In conclusion, the present study revealed various molecular responses to radiotherapy in rectal cancer. Cardiac muscle contraction may be affected by radiotherapy in patients with rectal cancer, while metastasis-associated biological processes may be prevented by radiotherapy. In addition, enriched protein domains such as the collagen triple helix repeat and olfactory receptors suggested that radiotherapy degrades the extracellular matrix and promotes metastasis in rectal cancer. In comparison with a previous study (12) that used the dataset GSE26027, the present study identified more DEGs using reasonable criteria, found a potential hub of DEGs by constructing a PPI network and identified TFs, oncogenes and TSGs in DEGs, providing information on the effects of radiotherapy at a molecular level. However, the hypotheses of the present study were formed based on the microarray data of only 6 patients. Therefore, additional exploration and validation is required, and the present authors are planning to use more samples to verify the DEGs with a high degree score in PPI network in future studies, and the annotated TFs, oncogenes and TSGs.


