Original Article
Identification of molecular target genes and key pathways in nasopharyngeal carcinoma by integrated bioinformatic analysis

Zhong-Guo Liang1*, Ye Li1*, Yu-Feng Lv1, Ding-Jian Liang1, Ren-Ba Liang1, Chun-Qiu Lin3, Yan-Xia Dai3, Xiao-Dong Zhu1

1Department of Radiation Oncology, The Affiliated Tumor Hospital of Guangxi Medical University, Cancer Institute of Guangxi Zhuang Autonomous Region, 71 Hedi Road, Nanning 530021, Guangxi, China; 2Department of Medical Oncology, Affiliated Langdong Hospital of Guangxi Medical University, 60 Xiangbin Road, Nanning 530021, Guangxi, China; 3Department of General Medicine, Guangxi Medical University, 22 Shuangyong Road, Nanning 530021, Guangxi, China. *Equal contributors and co-first authors.

Received November 30, 2018; Accepted January 8, 2019; Epub April 15, 2019; Published April 30, 2019

Abstract: Purpose: The current study aims to identify potent molecular target genes and key pathways involving the occurrence, development, and prognosis of nasopharyngeal carcinoma (NPC), and highlight their critical roles in management of NPC. Materials and methods: Expression profiling data for patients with NPC and non-cancer were searched and downloaded from the database of Gene Expression Omnibus. The data sets were applied to identify differentially expressed genes (DEGs) between the NPC group and the non-neoplastic group using the linear models for microarray analysis (limma) package in R language. Function enrichment analyses of DEGs was performed. The protein-protein interaction network was constructed and analyzed using String and Cytoscape, allowing the identification of hub genes for further analysis. Results: We identified 190 DEGs, of which 69 were upregulated and 121 genes were downregulated in NPC tissues compared to non-tumor tissues. Gene ontology enrichment analysis revealed that DEGs were mainly enriched in leukocyte migration, T cell activation, and cell adhesion. Kyoto Encyclopedia of Genes and Genomes pathway analysis indicated that the DEGs were mainly involved in nuclear factor-kappa B signalling pathway. Twenty hub genes were identified in the PPI network, including 5 up-regulated genes (FN1, PTGS2, CCND1, ITGAV, and STAT1), and 15 down-regulated genes (LCK, CD19, CCR7, ZAP70, RAC2, CD22, SELL, CD48, PTPN6, TNFRSF13C, BTK, CCL21, PLCG2, TNFSF11, and GNG7). Conclusion: Potential target genes and pathways were screened via integrated bioinformatic analysis. Future work is essential to verify the function of them across molecular biological experiments.

Keywords: Nasopharyngeal carcinoma, target gene, pathway, bioinformatic analysis

Introduction

Nasopharyngeal carcinoma (NPC) is prevalent in South-Eastern China, Malaysia, Indonesia, Singapore, Eastern Asia, and Northern Africa, with high incidences of 15-50/100,000 per year [1, 2]. According to a survey from the International Agency for Research on Cancer, there was an estimated 86,700 new cases and 50,800 related deaths in 2012 [2]. Patients with NPC are relatively asymptomatic at early stage as they originate in nasopharyngeal cavity, and therefore patients are often diagnosed with advanced disease when nodal metastasis occurs, or the tumor involves the adjacent critical normal structures like medial/lateral pterygoid, skull base, and the optic nerve. Though multidisciplinary approaches including radiotherapy, chemotherapy, and target therapy have been utilized, outcomes of these patients remain unsatisfactory, with relatively high rates of 5-year distant metastasis (15-21%) [3-5]. Therefore, it is worthwhile to identify other potential approaches to optimize survival for NPC.

Recently, microarray technology has become a high-throughput platform and an indispensable
Potent target genes and pathways for nasopharyngeal carcinoma

Figure 1. Identification of 190 commonly changed DEGs from the two cohort profile data sets (GSE40290, and GSE53819). Different color areas represent different datasets. The cross areas indicate the commonly changed DEGs. DEGs: differentially expressed genes.

Materials and methods

Microarray data and Identification of DEGs

Two gene expression profiles (GSE40290, and GSE53819) were downloaded from the database of GEO (https://www.ncbi.nlm.nih.gov/geo/). The array data of GSE40290 included 25 primary non-keratinizing NPCs and 8 nasopharyngitis tissues, based on the GPL8380 Capitalbio 22K Human oligo array version 1.0 platform (Capitalbio. Crop, Beijing, China). GSE53819 consisted of 18 NPC tissue samples and 18 non-cancerous nasopharyngeal tissues, based on the GPL6480 Agilent-014850 Whole Human Genome Microarray 4x44K G4112F platform (Agilent Technologies, Palo Alto, CA, USA). Given the genes corresponded to several probes, the average expression values of these probes were calculated to determine the expression value of the gene. Subsequently, the skewed distribution of data was converted into a normal distribution using a log 2 transformation, followed by normalization using the Median method [7]. The Linear Models for Microarray Analysis (limma) package [8] in R language was used to screen for the DEGs between the NPC and non-NPC tissue samples. |log2 Fold Change|>1 and P value <0.01 were set as the strict cut-offs for DEG identification.

Bidirectional hierarchical clustering was applied to DEGs based on Euclidean distance and displayed the results as a heat map.

Gene ontology and pathway enrichment

With the implementation of the R package of “clusterProfiler” methods [9], the functional profiles of gene and gene clusters were performed, including gene ontology (GO) [10] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment [11]. GO analyses included cell component (CC), biological process (BP) and molecular function (MF). A false discovery rate (FDR) <0.05 were considered significant.

Protein-protein interaction (PPI) network analysis

PPI network can help us obtain insights into interactions among DEGs, which correlate with oncogenesis, development and prognosis of NPC. The related information was gained by Search Tool for the Retrieval of Interacting Genes (STRING; http://string-db.org). An interaction with a combined score ≥0.4 was considered statistically significant. Then, a PPI network was constructed by Cytoscape software [12]. Moreover, cytoHubba, as a plugin of Cytoscape software [13], was employed to calculate...
the node degree, and the number of interconnections to screen the hub genes of PPI. The top twenty genes selected by degree algorithm [14] were defined as hub genes. Finally, the PPI network of hub genes was constructed by Cytoscape software [12].

Results

Identification of DEGs

In total, 803 and 1725 DEGs were respectively identified from GSE40290, and GSE53819
Potent target genes and pathways for nasopharyngeal carcinoma

datasets, with 190 genes screened out in both two datasets (Figure 1). Among the 190 genes, 69 were upregulated and 121 genes were downregulated in NPC tissues compared to non-tumor tissues (Table 1). In addition, the results of the expression analysis are presented as a heat map (Figures 2 and 3), and the results of hierarchical clustering showed these 190 DEGs expression patterns could basically distinguish NPC tissues and non-neoplastic tissues.

GO and KEGG pathway enrichment analysis

The annotation of the 190 DEGs was mainly classified into three functional groups: cellular component, biological process and molecular function groups (Figure 4A-C). Cellular component analysis revealed that the DEGs were significantly enriched in side of membrane, receptor complex, external side of plasma membrane, plasma membrane receptor complex, and cytoplasmic side of plasma membrane. In terms of biological process, the common DEGs are significantly enriched in leukocyte migration, T cell activation, leukocyte differentiation, leukocyte proliferation, and lymphocyte differentiation. Regarding molecular function, enriched GO terms were mainly cell adhesion molecule binding, sulphur compound binding, glycosaminoglycan binding, and heparin binding. In addition, by KEGG pathway analysis, six significantly enriched pathways were found, including nuclear factor (NF)-kappa B signalling pathway, B cell receptor signalling pathway, proteoglycans in cancer, hematopoietic cell lineage, small cell lung cancer, and primary immunodeficiency (Figure 4D).

Module selection from the PPI network

Based on the analysis in the STRING database, the PPI network of DEGs was constructed by Cytoscape software. As shown in Figure 5, the network consisted of 126 nodes and 277 edges. The pink nodes stand for up-regulated genes, and the green nodes represent down-regulated genes. In addition, twenty hub genes were filtered according to the degree algorithm. Five were up-regulated genes, including FN1,
Potent target genes and pathways for nasopharyngeal carcinoma

Figure 4. Significantly enriched GO annotation and KEGG pathways of differentially expressed genes between nasopharyngeal carcinoma and non-cancer groups. (A) Cellular component, (B) Biological process, (C) Molecular function. (D) KEGG pathway analysis.
Potent target genes and pathways for nasopharyngeal carcinoma

Figure 5. PPI Network of DEGs constructed by Cytoscape. The pink nodes stand for up-regulated genes, and the green nodes represent down-regulated genes.

Table 2. Twenty hub genes in network string_interactions.tsv ranked by Degree method

<table>
<thead>
<tr>
<th>Rank</th>
<th>Gene symbol</th>
<th>Gene description</th>
<th>Score</th>
<th>Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LCK</td>
<td>LCK proto-oncogene</td>
<td>23</td>
<td>Down</td>
</tr>
<tr>
<td>1</td>
<td>CD19</td>
<td>CD19 molecule</td>
<td>23</td>
<td>Down</td>
</tr>
<tr>
<td>3</td>
<td>FN1</td>
<td>Fibronectin 1</td>
<td>18</td>
<td>Up</td>
</tr>
<tr>
<td>4</td>
<td>CCR7</td>
<td>C-C motif chemokine receptor 7</td>
<td>17</td>
<td>Down</td>
</tr>
<tr>
<td>4</td>
<td>ZAP70</td>
<td>Zeta chain of T cell receptor associated protein kinase 70</td>
<td>17</td>
<td>Down</td>
</tr>
<tr>
<td>6</td>
<td>PTGS2</td>
<td>Prostaglandin-endoperoxide synthase 2</td>
<td>16</td>
<td>Up</td>
</tr>
<tr>
<td>7</td>
<td>RAC2</td>
<td>Rac family small GTPase 2</td>
<td>15</td>
<td>Down</td>
</tr>
<tr>
<td>8</td>
<td>CD22</td>
<td>CD22 molecule</td>
<td>12</td>
<td>Down</td>
</tr>
<tr>
<td>8</td>
<td>SELL</td>
<td>Selectin L</td>
<td>12</td>
<td>Down</td>
</tr>
<tr>
<td>8</td>
<td>CD48</td>
<td>CD48 molecule</td>
<td>12</td>
<td>Down</td>
</tr>
<tr>
<td>11</td>
<td>PTPN6</td>
<td>Protein tyrosine phosphatase, non-receptor type 6</td>
<td>11</td>
<td>Down</td>
</tr>
<tr>
<td>11</td>
<td>CCND1</td>
<td>Cyclin D1</td>
<td>11</td>
<td>Up</td>
</tr>
<tr>
<td>13</td>
<td>ITGAV</td>
<td>Integrin subunit alpha V</td>
<td>10</td>
<td>Up</td>
</tr>
<tr>
<td>13</td>
<td>TNFRSF13C</td>
<td>TNF receptor superfamily member 13C</td>
<td>10</td>
<td>Down</td>
</tr>
<tr>
<td>15</td>
<td>BTK</td>
<td>Bruton tyrosine kinase</td>
<td>9</td>
<td>Down</td>
</tr>
<tr>
<td>15</td>
<td>STAT1</td>
<td>Signal transducer and activator of transcription 1</td>
<td>9</td>
<td>Up</td>
</tr>
<tr>
<td>17</td>
<td>CCL21</td>
<td>C-C motif chemokine ligand 21</td>
<td>8</td>
<td>Down</td>
</tr>
<tr>
<td>17</td>
<td>PLCG2</td>
<td>Phospholipase C gamma 2</td>
<td>8</td>
<td>Down</td>
</tr>
<tr>
<td>17</td>
<td>TNFSF11</td>
<td>TNF superfamily member 11</td>
<td>8</td>
<td>Down</td>
</tr>
<tr>
<td>17</td>
<td>GNG7</td>
<td>G protein subunit gamma 7</td>
<td>8</td>
<td>Down</td>
</tr>
</tbody>
</table>

Figure 6. PPI Network of twenty hub genes constructed by Cytoscape.

PTGS2, CCND1, ITGAV, and STAT1. Fifteen were down-regulated genes, comprising of LCK, CD19, CCR7, ZAP70, RAC2, CD22, SELL, CD48, leukocyte migration, T cell activation, and cell adhesion, which were very associated with tumor growth, invasion, and metastasis [15,}

Discussion

In the current cohort, 190 DEGs were identified, including 69 upregulated and 121 downregulated genes between NPC and non-tumor tissues through significant analyses on microarray data. GO enrichment analysis revealed that DEGs were mainly enriched in
Potent target genes and pathways for nasopharyngeal carcinoma

Wang and colleagues observed that overexpression of epithelial cell adhesion molecule could result in epithelial-mesenchymal transition, stemness, and metastasis of NPC cells [17]. Additionally, targeting of focal adhesion proteins may potently sensitize cancer cells to radiotherapy, and chemotherapy [18].

Pathway enrichment indicated that the DEGs were mainly involved in NF-kappa B signalling pathway. Several studies have reported that this pathway played an important role in the occurrence, progression, and prognosis of several cancers, including prostate cancer [19], colon cancer [20], and head neck cancer [21]. It also correlates with EBV pathogenesis in NPC by regulating BamHI-A rightward transcripts [22]. When the pathway was activated, the radioresistance and chemoresistance could occur in NPC [23, 24]. Correspondingly, inhibition of the pathway might bring good anticancer results via inducing G2/M phase arrest and apoptosis [25]. Therefore, it may be a potential target pathway for improving therapeutic effects. More evidence is needed in this regard to confirm the efficacy of targeting this pathway in vivo and vitro.

By establishing a PPI, twenty hub genes were screened, which would provide new insights for NPC intervention strategy. FN1 is a member of the glycoprotein family, which involves in cell adhesion and migration processes [26, 27] that could promote cell proliferation and metastases by facilitating epithelial-mesenchymal transition (EMT) process of NPC cells [28]. Another mechanism has been reported that FN1 could inhibit apoptosis by upregulating BCL2, and promote migration and invasion [29]. In addition, expression of FN1 is negatively related to the prognosis of NPC including distant metastasis-free survival, and local recurrence-free survival [30]. Therefore, FN1 may be a potent therapeutic target for the treatment of NPC.

PTGS-2, also known as Cyclooxygenase-2 (COX-2), is the key enzyme in prostaglandin biosynthesis, and acts both as a dioxygenase and as a peroxidase. Overexpression of COX-2 was significantly found in patients with NPC. High level of its expression might contribute to a poor prognosis for NPC [31, 32]. First, high expression of COX-2 is related to high risk of lymph node metastasis [32], which has been proven to be an independent poor prognosis factor. Second, higher levels of COX-2 would result in increasing cell proliferation, and suppression of cellular senescence via the inactivation of p53, resulting in chemoresistance [31]. Moreover, it had been demonstrated that the expression of mitochondrial COX-2 would promote the stemness of NPC [33].

The protein encoded by CCND1 belongs to the highly conserved cyclin family. Wong and colleagues demonstrated that CCND1 was overexpressed in over 90% of NPC, and its activation played a critical role in NPC pathogenesis [34]. In addition, it might be an important target in regulating NPC cell cycle. It was observed that ribociclib, a specific cyclin dependent kinase (Cdk) 4/6 inhibitor, could lead to G1 arrest by blocking the formation of cyclin D1-Cdk4/6 complex [34]. It also appears that CCND1 is one of the crucial links in the process of regulating cell cycle for several micro-RNAs and long non-coding RNAs (IncRNAs), including miR-374a, miR-150, and IncRNA AK294004 [35-37]. For example, miR-374a could directly target CCND1 to inactivate pPI3K/pAKT/c-JUN forming a negative feedback loop, as well as suppressing downstream signals related to cell cycle progression and epithelial-mesenchymal transition (EMT), which would suppress NPC cell growth, metastasis and sensitizes NPC to cisplatin [35]. Additionally, IncRNA AK294004 could enhance radiation sensitivity via a negative effect on CCND1 [37]. Consequently, it seems promising to perform trials to investigate whether the target for CCND1 can translate to superior survival by improving radiosensitivity, and reducing metastasis for NPC.

The protein encoded by STAT1 is a member of the STAT protein family. Several evidences have shown Inhibition of STAT1 might help relieve immune tolerance in NPC by suppressing the expression of indoleamine 2,3-dioxygenase, which is a molecule of immune tolerance via decreasing T-cell proliferation. Furthermore, STAT1 may be closely associated with radiation resistance for NPC. Compared to CNE-2, significantly higher expression of STAT1 was reported in CNE-2R, a radioresistant cell line [38]. Inhibition of STAT1 could enhance radiosensitivity of CNE-2R by increasing the proportion of G2/M phase, suppressing growth, and promoting apoptosis in vitro and in vivo [39]. Hence, it
may be an optimal therapeutic target for radio-resistant NPC.

Various studies have shown that CD19 plays an important role in diagnosis and prognosis for NPC. First, the percentage of CD19\(^+\) circulating lymphocytes correlated negatively with TNM stage of NPC [40]. Secondly, after radiotherapy and chemotherapy, the percentage of CD19\(^+\) lymphocytes significantly decreased in NPC patients [40]. Moreover, patients with high circulating CD19\(^+\) B cell possessed significant better 5-year progression-free survival (PFS) than those with low circulating CD19\(^+\) B cell (81.8\% vs. 66.6\%, P=0.036) [41]. However, the mechanism is unclear why the level of CD19\(^+\) lymphocytes influences the long term survival of NPC.

PTPN6, also known as Src homology region 2 domain-containing phosphatase-1 (SHP-1). A dataset has shown that SHP-1 is negatively associated with radiation sensitivity for NPC. The mechanism may be that SHP-1 overexpression inhibits cellular senescence, enhance DNA DSB repair, increase S phase arrest and decrease cell apoptosis [42]. These were consistent with clinical observations by Peng et al. who observed high expression of SHP-1 was significantly associated with poor local recurrence-free survival [43] (P=0.008).

In conclusion, key target genes and pathways were screened via integrated bioinformatic analysis, which involved in the initiation, development, and progression of NPC. Given their potential roles in targeting NPC, future work is essential to verify the function of them across molecular biological experiments to enhance therapeutic outcomes.

Acknowledgements

This work was sponsored by the Basic Ability Enhancement Project of Young Teachers in Guangxi Zhuang Autonomous Region (2017-KY0114, 2017), the Youth Science Foundation of Guangxi Medical University (GXMUY-SF201515, 2015), and International Communication of Guangxi Medical University Graduate Education in 2018.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Xiao-Dong Zhu, The Affiliated Tumor Hospital of Guangxi Medical University, 71 Hedi Road, Nanning 530021, Guangxi, China. Tel: 86-771-5331466; E-mail: zhuxdonggxmu@126.com

References

Potent target genes and pathways for nasopharyngeal carcinoma


Potent target genes and pathways for nasopharyngeal carcinoma


