Abstract. Nasopharyngeal carcinoma (NPC) is one of the most common malignancies in the head and neck. The aim of the current study was to identify the key pathways and genes involved in NPC through bioinformatics analysis and to identify potential molecular mechanisms underlying NPC proliferation and progression. Three gene expression profiles (GSE12452, GSE34573 and GSE64634) were downloaded from the Gene Expression Omnibus database. A total of 76 samples were analyzed, of which 59 were NPC samples and 17 were normal samples. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were subsequently conducted. The protein-protein interaction (PPI) network of the differentially expressed genes (DEGs) was constructed using Cytoscape software. Analysis of GSE12452, GSE34573 and GSE64634 datasets identified 1,301 (553 upregulated and 748 downregulated), 1,232 (348 upregulated and 884 downregulated) and 1,218 (555 upregulated and 663 downregulated) DEGs, respectively. Using Venn diagram analysis, 268 DEGs (59 upregulated and 209 downregulated) that intersected all three datasets, were selected for further analysis. The results of GO analysis revealed that upregulated DEGs were significantly enriched in biological processes, including ‘cell adhesion’, ‘cell division’, ‘mitosis’ and ‘mitotic cell cycle’. The downregulated DEGs were mainly enriched in ‘microtubule-based movement’, ‘cilium movement’, ‘cilium axoneme assembly’ and ‘epithelial cell differentiation’. The KEGG pathway analysis results revealed that the upregulated DEGs were highly associated with several pathways, including ‘extracellular matrix-receptor interaction’, ‘human papillomavirus infection’, ‘arrhythmogenic right ventricular cardiomyopathy’ and ‘focal adhesion’, whereas the downregulated DEGs were enriched in ‘metabolic pathways’, ‘Huntington’s disease’, ‘fluid shear stress and atherosclerosis’ and ‘chemical carcinogenesis’. On the basis of the PPI network of the DEGs, the following top 10 hub genes were identified: Dynein axonemal light intermediate chain 1, dynein axonemal intermediate chain 2, calmodulin 1, coiled-coil domain containing 114, dynein axonemal heavy chain 5, radial spoke head 9 homolog, radial spoke head component 4A, NDC80 kinetochore complex component, thymidylate synthetase and coiled-coil domain containing 39. In conclusion, by performing a comprehensive bioinformatics analysis of DEGs, putative targets that could be used to elucidate the molecular mechanisms underlying NPC were identified.

Introduction

Nasopharyngeal carcinoma (NPC) is one of the most common malignancies in the head and neck (1). NPC has a unique geographical distribution, and is prevalent in Southeast Asia, the Middle East and North Africa (1). In endemic areas, NPC incidence may reach up to 35 cases per 100,000 persons among middle-aged males (2). The 5-year survival of patients with early-stage NPC is up to 95%, however, the survival rate of patients with advanced-stage NPC is only ~60% (3,4), and 70% of newly diagnosed patients with NPC have locoregionally advanced disease (5). Therefore, investigating potential biomarkers for the identification of patients with early-stage NPC is important to improve patient outcomes.
The presence of Epstein-Barr virus (EBV) DNA in plasma is currently used for screening asymptomatic patients with NPC, however, its positive predictive value for tumor screening is relatively low (11%) (6). Additionally, accumulating evidence indicates that polygenes and cell pathways, including the transforming growth factor-β signaling pathway and Notch signaling pathway, may contribute to the development and progression of NPC (7-9).

The precise molecular mechanisms underlying the progression of NPC remain unclear, and the early diagnosis and treatment of NPC is currently limited (10,11). Therefore, further studies to elucidate the molecular mechanisms involved in NPC proliferation and progression are required for a comprehensive understanding of NPC carcinogenesis.

Gene microarrays, which are high-throughput platforms for the analysis of gene expression, allow the identification of hundreds of differentially expressed genes (DEGs) involved in various signaling pathways, molecular functions and biological processes (12-14). However, only limited overlaps were observed when comparative analysis of the DEGs in independent studies was conducted (15,16). Combining microarray technologies and bioinformatics tools enhances the efficiency and accuracy of analysis (15,16). Wang et al (15) and Jiang et al (16) analyzed the GSE12452 dataset, which contained 31 NPC samples and 10 normal control samples, to identify the key genes involved in NPC. However, the number of samples included in these two studies was relatively small, and the molecular pathways involved in NPC carcinogenesis remain unclear. In the present study, the GSE12452 (17), GSE34573 (18) and GSE64634 (19) datasets were downloaded from the Gene Expression Omnibus database (GEO; www.ncbi.nlm.nih.gov/geo; GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array) to identify DEGs in NPC tissues. Subsequently, gene ontology (GO; www.geneontology.org) and pathway enrichment analysis were conducted to identify the biological functions and pathways of key genes (20). The results of the present study provides novel insights into potential biomarkers for NPC and may contribute to the current understanding of the molecular mechanisms underlying NPC proliferation and progression.

Materials and methods

Microarray data. Three gene expression profiles (GSE12452, GSE34573, and GSE64634) were downloaded from the GEO database. GSE12452, which was based on the Affymetrix GPL570 platform [GPL570 (HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array], was submitted by Ahlquist et al (17). The GSE12452 dataset contained 31 NPC samples and 10 normal NPC samples. The analysis for differential gene expression between tumor and normal tissue was performed using GeneSpring software version 11.5 (Agilent Technologies, Inc., Santa Clara, CA, USA). GSE34573, submitted by Hu et al (18), was based on the Affymetrix GPL570 platform [GPL570 (HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array] and consisted of 16 NPC samples and 3 normal control samples. GSE64634, submitted by Xiong et al, was based on the Affymetrix GPL570 platform [GPL570 (HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array] and consisted of 12 NPC samples and 4 normal controls (19). A Student’s t-test was used to identify DEGs with an alteration of ≥2-fold. P<0.05 was considered to indicate a statistically significant difference.

GO and pathway enrichment analysis of DEGs. GO analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG; www.genome.jp/kegg/pathway.html) pathway analysis were conducted to identify DEGs at the biologically functional level (21). The Database for Annotation, Visualization, and Integrated Discovery (DAVID; david.abcc.ncifcrf.gov) was used to integrate functional genomic annotations (22). P<0.05 was considered to indicate a statistically significant difference (23).

Integration of the protein-protein interaction (PPI) network. The Search Tool for the Retrieval of Interacting Genes version 10.0 (STRING; string-db.org) was used for the exploration of potential DEG interactions at the protein level (24). The PPI networks of DEGs by STRING were derived from validated experiments (25). A PPI score of >0.4 was considered significant. The PPI networks were visualized using Cytoscape software (http://www.cytoscape.org) (26). P<0.05 was considered to indicate a statistically significant difference.

Results

Identification of DEGs. NPC and normal samples (59 and 17, respectively) were first analyzed. GeneSpring software was used to analyze the series of each chip and to identify the DEGs. Following analysis of GSE12452, GSE34573, GSE64634 datasets, 1,301 (553 upregulated and 748 downregulated), 1,232 (348 upregulated and 884 downregulated) and 1,218 (555 upregulated and 663 downregulated) genes were identified, respectively. The results of the cluster analysis of DEGs revealed significant differences between the normal nasopharyngeal tissue and NPC samples (Fig. 1). Using Venn diagram analysis, 268 DEGs (59 upregulated and 209 downregulated) in the intersection of the above three datasets were selected for further analysis (Fig. 2).

GO term enrichment analysis. The identified DEGs were uploaded to the online software DAVID for GO and KEGG pathway analyses. The results of the GO analysis revealed that upregulated DEGs were significantly enriched in biological processes, including ‘cell adhesion’, ‘cell division’, ‘mitosis’, and ‘mitotic cell cycle’ (Table I; Fig. 3A). The downregulated DEGs were mainly enriched in ‘microtubule-based movement’, ‘cilium movement’, ‘cilium axoneme assembly’ and ‘epithelial cell differentiation’ (Table I; Fig. 3B). In terms of molecular function, the upregulated DEGs were enriched in ‘phosphatidylinositol-mediated signaling’, and the downregulated DEGs were enriched in ‘axonemal dynein complex assembly’ (Table I).

KEGG pathway analysis. KEGG pathway analysis revealed that the upregulated DEGs were highly associated with pathways including ‘ECM-receptor interaction’, ‘human papillomavirus infection’, ‘arrhythmogenic right ventricular cardiomyopathy’ and ‘focal adhesion’ (Table II; Fig. 4A). The downregulated DEGs were enriched in ‘metabolic pathways’,...
Figure 1. Cluster analysis of DEGs. The abscissa represents different samples; the vertical axis represents clusters of DEGs. Red color represents downregulation; green color represents upregulation. Expression data are represented as normalized values (Z-scores). (A) GSE12452 and (B) GSE34573.
‘Huntington’s disease’, ‘fluid shear stress’, ‘atherosclerosis’ and ‘chemical carcinogenesis’ (Table II; Fig. 4B).

PPI network. The DEG expression profiles in NPC were constructed according to the information in the STRING database. Following the elimination of isolated and partially connected nodes, a network of DEGs was constructed (Fig. 5). The top 10 hub genes, which were the genes exhibiting the most significant interaction, included dynein axonemal light intermediate chain 1 (DNALI1), dynein axonemal intermediate chain 2 (DNAI2), calmodulin 1 (CALM1), coiled-coil domain containing 114 (CCDC114), dynein axonemal heavy chain 5 (DNAH5), radial spoke head 9 homolog (RSPH9), radial spoke head component 4A (RSPH4A), NDC80 kinetochore complex component (NDC80), thymidylate synthetase (TYMS) and coiled-coil domain containing 39 (CCDC39). DNALI1 demonstrated the highest node degree of 18.

Discussion

NPC is one of the most common squamous cell tumors in the head and neck (1). The 5-year survival rate among
patients with stage I disease is 95% (3). However, the 5-year survival rate among patients with stage IV disease is just over 60% (27). Therefore, understanding the etiological factors and molecular mechanisms of NPC progression is essential for diagnosis and treatment. Microarray technology has been widely applied to predict the potential therapeutic targets for carcinoma, including colorectal cancer (12-14). Previously, Wang et al (15) analyzed the GSE12452 dataset and revealed that cyclin B1, mitotic arrest deficient 2 like 1, proliferating cell nuclear antigen, mucin 1, cell surface associated and aldehyde dehydrogenase 1 family member A1 may be involved in EBV-associated NPC (15). A study analyzing the GSE12452
Table I. Gene ontology analysis of differentially expressed genes associated with nasopharyngeal carcinoma.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Term/gene function</th>
<th>Gene count</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN1, ITGA5, ADAM23, DSG2, SSX2IP, COL5A1, LAMB1, CNTNAP2</td>
<td>GO:0007155/cell adhesion</td>
<td>8</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>CCND2, HELLS, AURKA, SMC4, NDC80, HAUS6</td>
<td>GO:0051301/cell division</td>
<td>6</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>HAUS6, NDC80, ANLN, AURKA, HELLS</td>
<td>GO:0007067/mitosis</td>
<td>5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>SMC4, RFC4, NUP107, NDC80, TYMS, AURKA</td>
<td>GO:000278/mitotic cell cycle</td>
<td>6</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>FN1, COL5A1, COL22A1, ITGA5, COL4A1</td>
<td>GO:0030198/extracellular matrix organization</td>
<td>5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>LAMB1, FZD7, FN1</td>
<td>GO:0034446/substrate adhesion-dependent cell spreading</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>CXCL10, NTRK2</td>
<td>GO:0010996/response to auditory stimulus</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>RFC4, NDC80, AURKA, TYMS</td>
<td>GO:0048015/phosphatidylinositol-mediated signaling</td>
<td>4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>COL4A1, LHX2, LAMB1, ITGA5, COL5A1</td>
<td>GO:0007411/axon guidance</td>
<td>5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>TYMS, MTHFD2</td>
<td>GO:0046653/tetrahydrofolate metabolic process</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>TYMS, HOX3, PRRX1</td>
<td>GO:0051216/cartilage development</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>FN1, COL4A1, ITGA5, HOX3</td>
<td>GO:0001525/angiogenesis</td>
<td>4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>NTRK2, ARNT2, ITGA5, HOX3, CXCL10</td>
<td>GO:0008284/positive regulation of cell proliferation</td>
<td>5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>RIF1, NDC80, CCND2, AURKA</td>
<td>GO:0007049/cell cycle</td>
<td>4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>NUP107, SMC4, NDC80</td>
<td>GO:0000236/mitotic prometaphase</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>SMC4, NDC80</td>
<td>GO:0000070/mitotic sister chromatid segregation</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AURKA, NDC80</td>
<td>GO:0007051/spindle organization</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>FZD7, GRB10</td>
<td>GO:0042327/positive regulation of phosphorylation</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>FN1, LAMB1, COL5A1</td>
<td>GO:0016477/cell migration</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>RFC4, DTL, TYMS</td>
<td>GO:0006260/DNA replication</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>CNTNAP2, NTRK2</td>
<td>GO:0021987/cerebral cortex development</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>GJA1, OLFM1</td>
<td>GO:0051259/protein oligomerization</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DMD, CNTNAP2</td>
<td>GO:0034613/cellular protein localization</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>NTRK2</td>
<td>GO:0048935/peripheral nervous system neuron development</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>NTRK2</td>
<td>GO:0050773/regulation of dendrite development</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>DMD</td>
<td>GO:0060314/regulation of ryanodine-sensitive calcium-release channel activity</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>RIF1, DTL</td>
<td>GO:0006974/response to DNA damage stimulus</td>
<td>2</td>
<td>0.04</td>
</tr>
<tr>
<td>NTRK2, AURKA</td>
<td>GO:0046777/protein autophosphorylation</td>
<td>2</td>
<td>0.04</td>
</tr>
<tr>
<td>SRSF10</td>
<td>GO:0006376/mRNA splice site selection</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>FZD7</td>
<td>GO:0010812/negative regulation of cell-substrate adhesion</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>TFC</td>
<td>GO:0045780/positive regulation of bone resorption</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>GJA1</td>
<td>GO:0060174/limb bud formation</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>LHX2</td>
<td>GO:2000678/negative regulation of transcription regulatory region DNA binding</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>NUP107, NDC80</td>
<td>GO:000090/mimetic anaphase</td>
<td>2</td>
<td>0.05</td>
</tr>
<tr>
<td>HELLS</td>
<td>GO:001655/urogenital system development</td>
<td>1</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Table I. Continued.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Term/gene function</th>
<th>Gene count</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNLY</td>
<td>GO:0031640/killing of cells of other organism</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>CXCL10</td>
<td>GO:0033280/response to vitamin D</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>SRSF10</td>
<td>GO:0048025/negative regulation of mRNA splicing, via spliceosome</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>DSG2</td>
<td>GO:0086005/regulation of ventricular cardiac muscle cell action potential</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>DNAAF1, DNAAF3, DRC1, CCDC39</td>
<td>GO:0070286/axonemal dynein complex assembly</td>
<td>4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DNH12, DNAH5, DNAH9, DYNLRB2, DYNC2H1, DNAH10, KIF9</td>
<td>GO:0007018/microtubule-based movement</td>
<td>7</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>RSPH4A, RSPH9, DNAAF1</td>
<td>GO:0003341/cilium movement</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AK7, RSPH4A, RSPH9</td>
<td>GO:001539/flagsellar motility</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AKR1C2, AKR1C3, DHR9</td>
<td>GO:0042448/progesterone metabolic process</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>MUC4, MUC16, MUC5AC, MUC1, MUC20</td>
<td>GO:0016266/O-glycan processing</td>
<td>5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>ELF3, DHR9, EZR, UPK1B</td>
<td>GO:0030855/epithelial cell differentiation</td>
<td>4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AKR1C3, AKR1C2</td>
<td>GO:0034694/response to prostaglandin stimulus</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>ANKRD45, CFAP53, SORBS2, SCGB2A1, TMEM125, DNAH5, MUC4, MSMB, RSPH14, LRRC34, VNN3</td>
<td>GO:0008150/biological_process</td>
<td>11</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DNAAF1, DNAAF3</td>
<td>GO:0044458/motile cilium assembly</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DYNCH1, ROPN1L, DNAH5</td>
<td>GO:0001539/flagsellar motility</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>RDH10, DHR9</td>
<td>GO:0002138/retinoid acid biosynthetic process</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DNAAF1, CCDC14</td>
<td>GO:0036158/outer dynein arm assembly</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DRC1, CCDC39</td>
<td>GO:0060286/flagsellar cell motility</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AKR1C2, AKR1C3</td>
<td>GO:0071395/cellular response to jasmonic acid stimulus</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>GPRC5A, CDS1, CALM1, RIIA1, CCL15, EPAS1, CHL1, RRAD, ROPN1L, CEACAM6, GNAL, SCGB1A1, CAPN5</td>
<td>GO:0007165/signal transduction</td>
<td>13</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>SERPINB4, WFD2C, SLPI, SERPINB3</td>
<td>GO:0010951/negative regulation of endopeptidase activity</td>
<td>4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>CHST6, CKB, ADH1C, CALM1, AKR1C3, SGMS2, ASS1, NQO1, SLC44A4, CHST9, CDS1, AK1, DHR24, CH25H, GSTA1</td>
<td>GO:0044281/small molecule metabolic process</td>
<td>15</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>SPAG6, DNAH9, PIFO</td>
<td>GO:0030030/cell projection organization</td>
<td>3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DYNCH1, DNA12, B9D1, CC2D2A</td>
<td>GO:0042384/cilium assembly</td>
<td>4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>CCDC39, DNAAF1</td>
<td>GO:0035469/determination of pancreatic left/right asymmetry</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DRC1, CCDC39</td>
<td>GO:0060286/flagsellar cell motility</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>CCDC39, DNAAF1</td>
<td>GO:0071907/determination of digestive tract left/right asymmetry</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>CCDC39, DNAAF1</td>
<td>GO:0071910/determination of liver left/right asymmetry</td>
<td>2</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

DEGs, differentially expressed genes.
Table II. Kyoto Encyclopedia Analysis of Genes and Genomes pathway analysis of DEGs associated with nasopharyngeal carcinoma.

<table>
<thead>
<tr>
<th>Pathway ID</th>
<th>Name</th>
<th>Gene count</th>
<th>P-value</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>04512</td>
<td>ECM-receptor interaction</td>
<td>5</td>
<td>&lt;0.01</td>
<td>ITGA V, COL4A1, FN1, HMMR, LAMB1</td>
</tr>
<tr>
<td>05165</td>
<td>Human papillomavirus infection</td>
<td>6</td>
<td>&lt;0.01</td>
<td>FZD7, ITGA V, LAMB1, CCND2, COL4A1, FN1</td>
</tr>
<tr>
<td>05412</td>
<td>Arrhythmogenic right ventricular cardiomyopathy</td>
<td>4</td>
<td>&lt;0.01</td>
<td>DMD, GJA1, ITGA V, DSG2</td>
</tr>
<tr>
<td>04510</td>
<td>Focal adhesion</td>
<td>5</td>
<td>&lt;0.01</td>
<td>LAMB1, FN1, ITGA V, CCND2, COL4A1</td>
</tr>
<tr>
<td>05222</td>
<td>Small cell lung cancer</td>
<td>4</td>
<td>&lt;0.01</td>
<td>ITGA V, FN1, COL4A1, LAMB1</td>
</tr>
<tr>
<td>05200</td>
<td>Pathways in cancer</td>
<td>6</td>
<td>&lt;0.01</td>
<td>ARNT2, FN1, LAMB1, FZD7, COL4A1, ITGA V</td>
</tr>
<tr>
<td>04151</td>
<td>PI3K-Akt signaling pathway</td>
<td>5</td>
<td>&lt;0.01</td>
<td>FN1, COL4A1, CCND2, ITGA V, LAMB1</td>
</tr>
<tr>
<td>04974</td>
<td>Protein digestion and absorption</td>
<td>3</td>
<td>&lt;0.01</td>
<td>COL22A1, COL4A1, COL5A1</td>
</tr>
<tr>
<td>05146</td>
<td>Amoebiasis</td>
<td>3</td>
<td>&lt;0.01</td>
<td>LAMB1, COL4A1, FN1</td>
</tr>
<tr>
<td>00670</td>
<td>One carbon pool by folate</td>
<td>2</td>
<td>&lt;0.01</td>
<td>MTHFD2, TYMS</td>
</tr>
<tr>
<td>00250</td>
<td>Alanine, aspartate and glutamate metabolism</td>
<td>2</td>
<td>&lt;0.01</td>
<td>GAD1, ASNS</td>
</tr>
<tr>
<td>05205</td>
<td>Proteoglycans in cancer</td>
<td>3</td>
<td>&lt;0.01</td>
<td>FZD7, FN1, ITGA V</td>
</tr>
<tr>
<td>05410</td>
<td>Hypertrophic cardiomyopathy</td>
<td>2</td>
<td>0.01</td>
<td>ITGA V, DMD</td>
</tr>
<tr>
<td>05414</td>
<td>Dilated cardiomyopathy</td>
<td>2</td>
<td>0.01</td>
<td>DMD, ITGA V</td>
</tr>
<tr>
<td>04933</td>
<td>AGE-RAGE signaling pathway in diabetic complications</td>
<td>2</td>
<td>0.02</td>
<td>FN1, COL4A1</td>
</tr>
<tr>
<td>04620</td>
<td>Toll-like receptor signaling pathway</td>
<td>2</td>
<td>0.02</td>
<td>CXCL11, CXCL10</td>
</tr>
<tr>
<td>03040</td>
<td>Spliceosome</td>
<td>2</td>
<td>0.03</td>
<td>U2SURP, SRSF10</td>
</tr>
<tr>
<td>04550</td>
<td>Signaling pathways regulating pluripotency of stem cells</td>
<td>2</td>
<td>0.03</td>
<td>FZD7, RIF1</td>
</tr>
<tr>
<td>04310</td>
<td>Wnt signaling pathway</td>
<td>2</td>
<td>0.04</td>
<td>FZD7, CCND2</td>
</tr>
<tr>
<td>04514</td>
<td>Cell adhesion molecules (CAMs)</td>
<td>2</td>
<td>0.04</td>
<td>ITGA V, CNTNAP2</td>
</tr>
<tr>
<td>04150</td>
<td>mTOR signaling pathway</td>
<td>2</td>
<td>0.04</td>
<td>GRB10, FZD7</td>
</tr>
<tr>
<td>04145</td>
<td>Phagosome</td>
<td>2</td>
<td>0.04</td>
<td>ITGA V, TFRC</td>
</tr>
<tr>
<td>04390</td>
<td>Hippo signaling pathway</td>
<td>2</td>
<td>0.04</td>
<td>FZD7, CCND2</td>
</tr>
<tr>
<td>00430</td>
<td>Taurine and hypotaurine metabolism</td>
<td>1</td>
<td>0.04</td>
<td>GAD1</td>
</tr>
<tr>
<td>01100</td>
<td>Metabolic pathways</td>
<td>16</td>
<td>&lt;0.01</td>
<td>DHR59, CDS1, CKB, AK7, ADH1C, RDH10, SGMS2, DHCR24, AK1, ALDH3A1, AK9, ALDH3A2, ASS1, ST6GALNAC1, CHST9, AKR1C3</td>
</tr>
<tr>
<td>05016</td>
<td>Huntington’s disease</td>
<td>6</td>
<td>&lt;0.01</td>
<td>DNALI1, DNA2, DNAH10, DNAH12, DNAH9, DNAH5</td>
</tr>
<tr>
<td>05418</td>
<td>Fluid shear stress and atherosclerosis</td>
<td>5</td>
<td>&lt;0.01</td>
<td>CALM1, NQO1, CALML4, ASS1, GSTA1</td>
</tr>
<tr>
<td>05204</td>
<td>Chemical carcinogenesis</td>
<td>4</td>
<td>&lt;0.01</td>
<td>ALDH3A1, AKR1C2, ADH1C, GSTA1</td>
</tr>
<tr>
<td>00730</td>
<td>Thiamine metabolism</td>
<td>2</td>
<td>0.01</td>
<td>AK7, AK1</td>
</tr>
<tr>
<td>00830</td>
<td>Retinol metabolism</td>
<td>3</td>
<td>0.01</td>
<td>DHR59, RDH10, ADH1C</td>
</tr>
<tr>
<td>00010</td>
<td>Glycolysis/Gluconeogenesis</td>
<td>3</td>
<td>0.01</td>
<td>ALDH3A1, ALDH3A2, ADH1C</td>
</tr>
<tr>
<td>00982</td>
<td>Drug metabolism-cytochrome P450</td>
<td>3</td>
<td>0.01</td>
<td>GSTA1, ALDH3A1, ADH1C</td>
</tr>
<tr>
<td>00980</td>
<td>Metabolism of xenobiotics by cytochrome P450</td>
<td>3</td>
<td>0.01</td>
<td>ALDH3A1, ADH1C, GSTA1</td>
</tr>
</tbody>
</table>
dataset suggested that C-X-C motif chemokine ligand (CXCL) 9, ZIC family member 2, prostaglandin-endoperoxide synthase 2, fibronectin 1, CXCL10 and ovo like transcriptional repressor 1 may serve roles in NPC (16). However, the number of samples from individual datasets was relatively small (15,16). In the current study, 3 datasets were analyzed and 53 upregulated and 209 downregulated DEGs were screened by bioinformatics analysis.

The results of the KEGG pathway enrichment analysis and GO function annotation revealed that upregulated DEGs were mainly enriched in ‘cell adhesion’, ‘cell division’, ‘mitosis’, ‘mitotic cell cycle’, ‘ECM-receptor interaction’ and ‘human papillomavirus infection’, whereas downregulated DEGs were mainly involved in ‘axonemal dynein complex assembly’, ‘microtubule-based movement’, ‘metabolic pathways’, ‘Huntington’s disease’, ‘fluid shear stress’ and ‘atherosclerosis’ and ‘chemical carcinogenesis’. Previous studies have demonstrated that upregulation or downregulation of specific genes may affect NPC cell invasion, metastasis, proliferation and apoptosis (28-30). This result is consistent with the fact that carcinoma cell invasion and metastasis are closely associated with abnormal cell adhesion and cell division (28-30). Furthermore, cancer cell proliferation and apoptosis are closely associated with abnormalities in the mitotic cell cycle (28-30).

In the present study, a PPI network was constructed to identify the 10 most significant hub genes. These were as follows: DNALI1, DNAI2, CALM1, CCDC114, DNAH5, RSPH9, RSPH4A, NDC80, TYMS, and CCDC39. DNALI1 was the hub gene exhibiting the highest degree of connectivity. Peng et al (33) revealed that the mRNA levels of DNALI1 were significantly reduced in patients with allergic nasal mucosa compared with controls (P<0.05). Parris et al (34) reported that several malignant tumors with normal gene dosage levels displayed DNALI1 downregulation, suggesting that DNALI1 may be a novel therapeutic target for cancer drug development. The second hub gene identified, DNAI2, which is also protein encoding, is associated with primary ciliary dyskinesia (PCD) (35). DNAI2 and forkhead box J1 are ciliated cell markers (36). The third hub gene, CALM1, is one of genes encoding the calmodulin protein (37). Kim et al (38) conducted large-scale genome analyses for breast cancer, and the results indicated that, as a potential regulator of protein kinase B, CALM1 was highly expressed in phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α-mutated breast cancer. Furthermore, calcium-binding proteins CALM1, calumenin and reticulocalbin 1, were significantly upregulated in irradiated tumor cells, which were subjected to hypoxia, indicating that these mediators serve important roles in promoting tumor cell survival during hypoxia (39). Similar to DNAI2, CCDC114 is one of the PCD-associated genes, in which loss-of-function mutations result in PCD with laterality...
malformations involving heart defects (40). The absence or mislocalization of another hub gene, DNAH5, is a characteristic marker for motile ciliary abnormality in nasal polyps (41). The remaining five hub genes in the present study were RSPH9, RSPH4A, NDC80, TYMS and CCDC39. Yoon et al (42) reported that the RSPH9 methylation pattern is a prognostic indicator in patients with nonmuscle invasive bladder cancer. RSPH9 and RSPH4A are radial spoke head protein genes, wherein mutations cause primary ciliary dyskinesia with central-microtubular-pair abnormalities (43). TYMS is a key enzyme in the de novo synthesis of 2'-deoxythymidine-5'-monophosphate from 2'-deoxyuridine-5'-monophosphate (44). CCDC39 and CCDC40 were first identified as causative mutations in patients with primary ciliary dyskinesia and are likely to be involved in the recruitment of tubulin glutamylase(s) to the flagella (45), which has not been identified to be associated with the development of NPC.

In conclusion, the present study conducted a comprehensive bioinformatics analysis of DEGs which may be involved in NPC progression. The results may provide novel insights into targets that can be used for the future investigation of molecular mechanisms underlying NPC. However, the specific functions of the identified genes in NPC should be confirmed by further molecular biological experiments.

Acknowledgements
Not applicable.

Funding
No funding was received.

Availability of data and materials
The datasets used during the present study are available from the corresponding author upon reasonable request.

Authors’ contributions
HMZ and QF conceived and designed the study. HMZ, QF, LXQ, BLL, LY and XH performed the bioinformatics analysis. LXQ and BLL analyzed the data. HMZ and QF wrote the manuscript. LY and XH reviewed and checked the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.
Clinical Analysis of HOPX hypermethylation promotes Upregulated long non-coding RNA

Chen X, Zhang T, He Q, Li B, Li Y, et al.


Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References


