Original Article

Variations in adamantinomatous craniopharyngioma: a gene set enrichment analysis

Yangfan Zou1,6#, Lin Wu1#, Yuanmeng Zhang2, Xin Zeng3, Zhaokai He4, Zhantao Jiang1, Xiuxiu Chen1, Lingbing Meng5*, Xin Yu1,6*

1Department of Neurosurgery, Chinese PLA General Hospital-Sixth Medical Center, No. 6, Fucheng Road, Haidian District, Beijing 100037, P. R. China; 2Department of Cardiology, The Third Medical Centre of Chinese PLA General Hospital, Yongding Road, Hai Dian, Beijing 100039, P. R. China; 3Department of Clinical Medicine, Southwest Medical University, Luzhou 646000, Sichuan Province, P. R. China; 4National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, State Key Laboratory for Infectious Disease Prevention and Control, No. 5 Liuzi, Changping, Beijing 102200, P. R. China; 5Neurology Department, Beijing Hospital, National Center of Gerontology, No. 1 Dahua Road, Dong Dan, Beijing 100730, P. R. China; 6Department of Neurosurgery, Affiliated Navy Clinical College of Anhui Medical University, Beijing 100037, P. R. China.

*Equal contributors. #Equal contributors.

Received May 25, 2020; Accepted September 17, 2020; Epub December 15, 2020; Published December 30, 2020

Abstract: Adamantinomatous craniopharyngioma (ACP) is an epithelial tumor that occurs in the sellar region; it may be derived from embryonic residue of the Rathke fissure epithelium. Surgery is difficult to perform and resection success rates are low because the tumor can invade and compress important surrounding structures. Therefore, it is necessary to find a safer and more effective treatment for ACP. Our research involved access to microarray data (GSE68015), and we identified differentially expressed genes (DEGs) between ACP and healthy samples using the limma package. We performed a gene set enrichment analysis (GSEA) based on Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG), constructed and analyzed a protein-protein interaction (PPI) network, and identified the significant modules. We identified 75 DEGs, comprising 64 upregulated genes and 11 downregulated genes in ACP samples compared with their expressions in healthy brain samples. The GSEA of these DEGs provided a comprehensive overview of some major pathophysiological mechanisms involved in ACP, including the roles played by the intermediate filaments, transport vesicles, and membranes. We also built a PPI network and identified five key genes: CDH1, KRT19, KRT16, KRT5, and KRT14. When comparing the ACP and healthy brain samples, 75 DEGs and five hub genes were identified that may be involved in the occurrence and progression of ACP, in particular the genes involved with the intermediate filaments.

Keywords: Adamantinomatous craniopharyngioma, gene set enrichment analysis, intermediate filament, gene expression omnibus

Introduction

Adamantinomatous craniopharyngioma (ACP) is an epithelial tumor that occurs in the sellar region [1, 2]. ACP can affect patients’ vision and visual field and is caused by the tumor’s compression of the optic nerve [3, 4]. Some patients may also suffer from increased intracranial pressure, hydrocephalus, and consciousness disorders [5]. However, this tumor’s propensity to invade or compress the surrounding important structures makes surgery very difficult, so its resection success rate is low [6]. Furthermore, ACP can develop into squamous cell carcinoma following radiotherapy [7]. Therefore, it is necessary to find a safer and more effective treatment for ACP.

With developments in molecular biology and technologies for targeted gene therapy, an increasing number of studies of the molecular pathology of ACP have recently been conducted [8-11]. Studies have confirmed that the over-activation of the WNT/β-Catenin signaling pathway is the molecular basis of ACP [12-15]. However, due to the lack of an effective diagnosis and treatment for ACP in its early stages, its early diagnosis and treatment remain difficult
Analysis of adamantinomatous craniopharyngioma

[16, 17]. Therefore, in order to develop an effective diagnosis and treatment for ACP, it is vital that we ascertain the molecular mechanisms underlying its occurrence, proliferation, and recurrence.

Bioinformatics can help us to identify differentially expressed genes (DEGs) and functional pathways related to the occurrence and development of tumors, such as ACP [18-20].

In this study, we analyzed data downloaded from the Gene Expression Omnibus (GEO) database to screen DEGs between samples of brain tissue from healthy individuals and samples from patients with ACP, in order to explore the potentially important molecular biological mechanisms.

Materials and methods

Access to microarray data

A gene expression profile, GSE68015, was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) [21]. The GSE68015 array data comprised mRNA expression profiles of 15 samples from patients with ACP and 16 samples from healthy individuals. These gene expression profiles were sequenced using Affymetrix HG-U133plus2 chips (Platform GPL570). All the ACP samples were collected from craniopharyngioma surgery tissues, and the control samples were obtained from the brain tissues of healthy individuals without ACP.

Identification of DEGs

The linear models for microarray data (limma) package is a full-featured package that contains both a RAW data input, a pre-processing (normalization) function for cDNA chips, and limma for the analysis of DEGs, especially for multifactor-design experiments. The limma package was used to identify the DEGs between the ACP and healthy brain tissue samples. The cut-off criteria were set as an adjusted p-value < 0.001 and a |logFC| ≥ 7.

GO and KEGG gene set enrichment analysis

Gene Ontology (GO) includes three aspects of biology: cellular components, molecular functions, and biological processes [22]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.kegg.jp/) aims to ascertain advanced functions and biological systems [23]. Gene set enrichment analysis (GSEA) can be used to analyze all sequenced genes in two sample groups. Its input is a gene expression matrix in which the genes are divided into two groups; all the genes are first sequenced and then used to indicate changing trends in the expression of the genes between the two groups. GSEA analyzes whether all the genes of an a priori defined set are enriched at the top or bottom of this sequence list [24]. GO and KEGG pathway enrichment analysis was performed for the identified DEGs using GSEA. GSEA was also conducted for all the sequenced genes from the ACP tumor tissues and the healthy brain tissues, using GSEA software, after importing the gene annotation files, reference function sets, and all gene data from both the ACP tumor tissues and the healthy brain tissues. This software performs analyses and sequences genes according to an algorithm, to produce a gene sequence list. It then analyzes the position of all the genes in the sequence list and scores them to obtain an enrichment score (ES). Once the ES has been standardized, we can obtain a comprehensive understanding of the biological functions of the genes through the enrichment of functional sets. The cut-off criterion is P < 0.05.

Construction and analysis of the protein-protein interaction (PPI) network and significant modules

The common DEGs were imported into the Search Tool for the Retrieval of Interacting Genes (STRING; http://string-db.org) (version 10.5), an online tool used to predict and trace-out PPI networks [25]. Cytoscape (version 3.6.1), a free visualization software application, was used to visualize the PPI networks [26]. Based on topology principles, the molecular complex detection (MCODE) (version 1.5.1) plug-in of Cytoscape can be used to discover tightly coupled regions [27]. MCODE was used to screen the modules of the PPI network with a degree cut-off of 6 and a node score cut-off of 0.2. Once the degrees were set (degrees 10), the hub genes were excavated.

Statistical analysis

SPSS software (version 25.0; IBM, Chicago, IL, USA) was used to perform the statistical analy-
Table 1. Functional enrichment analysis of the DEGs in ACP, using GSEA

<table>
<thead>
<tr>
<th>Gene set name</th>
<th>SIZE</th>
<th>ES</th>
<th>NES</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upregulated</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO_REGULATION_OF_TRANSCRIPTION_ELONGATION_FROM_RNA_POLYMERASE_II_PROMOTER</td>
<td>22</td>
<td>0.68</td>
<td>2.13</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_INTERMEDIATE_FILEMEMENT_BASED_PROCESS</td>
<td>32</td>
<td>0.69</td>
<td>1.91</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_REGULATION_OF_ACTIVIN_RECEPTOR_SIGNALING_PATHWAY</td>
<td>23</td>
<td>0.66</td>
<td>1.82</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_LATERAL_PLASMA_MEMBRANE</td>
<td>44</td>
<td>0.63</td>
<td>1.81</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_REGULATION_OF_DNA_BINDING</td>
<td>82</td>
<td>0.59</td>
<td>1.80</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_REGULATION_OF_MITOCHONDRIAL_OUTER_MEMBRANE_PERMEABILIZATION_INVOLVED_IN_APOPTOTIC_SIGNALING_PATHWAY</td>
<td>37</td>
<td>0.66</td>
<td>1.79</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Downregulated</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO_GLUTAMATE_RECEPTOR_BINDING</td>
<td>34</td>
<td>-0.75</td>
<td>-1.70</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_TRANSPORT_VESICLE_MEMBRANE</td>
<td>128</td>
<td>-0.61</td>
<td>-1.68</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_REGULATION_OF_RESPONSE_TO_FOOD</td>
<td>16</td>
<td>-0.74</td>
<td>-1.66</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_EXOCYTIC_VESICLE_MEMBRANE</td>
<td>49</td>
<td>-0.85</td>
<td>-1.63</td>
<td>0.000</td>
</tr>
<tr>
<td>GO_VESICLE_DOCKING_INVOLVED_IN_EXOCYTOSIS</td>
<td>32</td>
<td>-0.56</td>
<td>-1.61</td>
<td>0.017</td>
</tr>
<tr>
<td>GO_ADULT_BEHAVIOR</td>
<td>110</td>
<td>-0.69</td>
<td>-1.60</td>
<td>0.000</td>
</tr>
</tbody>
</table>

DEGs: differentially expressed genes; ACP: adamantinomatous craniopharyngioma; ES: enrichment score; NES: normalized enrichment score.
Figure 1. Six significant enrichment plots from the functional enrichment analyses of the DEGs in ACP, using GSEA.
A. GO_REGULATION_OF_TRANSCRIPTION_ELONGATION_FROM_RNA_POLYMERASE_II_PROMOTER. B. GOINTERMEDIATE_FILAMENT_BASED_PROCESS. C. GO_REGULATION_OF_AKTIVIN_RECEPTOR_SIGNALING_PATHWAY. D. GO_GLUTAMATE_RECEPTOR_BINDING. E. GO_REGULATION_OF_RESPONSE_TO_FOOD. F. GO_VESICLE_DOCKING_INVOLVED_IN_EXOCYTOSIS.
A. GO_INTERMEDIATE_FILAMENT_BASED_PROCESS. C. GO_REGULATION_OF_ACTIVIN_RECEPTOR_SIGNALING_PATHWAY. D. GO_GLUTAMATE_RECEPTOR_BINDING. E. GO_REGULATION_OF_RESPONSE_TO_FOOD. F. GO_VESICLE_DOCKING_INVOLVED_IN_EXOCYTOSIS.

Table 2. Pathway enrichment analysis of the DEGs in ACP, using GSEA

<table>
<thead>
<tr>
<th>Gene set name</th>
<th>SIZE</th>
<th>ES</th>
<th>NES</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upregulated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEGG_PATHWAYS_IN_CANCER</td>
<td>287</td>
<td>0.56</td>
<td>1.67</td>
<td>0.000</td>
</tr>
<tr>
<td>KEGG_BLADDER_CANCER</td>
<td>37</td>
<td>0.69</td>
<td>1.67</td>
<td>0.002</td>
</tr>
<tr>
<td>KEGG_CHRONIC_MYEOID_LEUKEMA</td>
<td>63</td>
<td>0.59</td>
<td>1.66</td>
<td>0.006</td>
</tr>
<tr>
<td>KEGG_SMALL_CELL_LUNG_CANCER</td>
<td>77</td>
<td>0.64</td>
<td>1.63</td>
<td>0.000</td>
</tr>
<tr>
<td>KEGG_PATHOGENIC_ESCHERICHIA_COLI_INFECTION</td>
<td>42</td>
<td>0.55</td>
<td>1.60</td>
<td>0.025</td>
</tr>
<tr>
<td>KEGG_MELANOMA</td>
<td>64</td>
<td>0.53</td>
<td>1.59</td>
<td>0.006</td>
</tr>
<tr>
<td>Downregulated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEGG_PROXIMAL_TUBULE_BICARBONATE_RECLAMATION</td>
<td>23</td>
<td>-0.55</td>
<td>-1.44</td>
<td>0.036</td>
</tr>
<tr>
<td>KEGG_CARDIAC_MUSCLE_CONTRACTION</td>
<td>62</td>
<td>-0.46</td>
<td>-1.37</td>
<td>0.103</td>
</tr>
<tr>
<td>KEGG_ALANINEASPARTATE_AND_GLUTAMATE_METABOLISM</td>
<td>30</td>
<td>-0.49</td>
<td>-1.36</td>
<td>0.096</td>
</tr>
<tr>
<td>KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION</td>
<td>203</td>
<td>-0.47</td>
<td>-1.30</td>
<td>0.080</td>
</tr>
</tbody>
</table>

DEGs: differentially expressed genes; ACP: adamantinomatous craniopharyngioma; ES: enrichment score; NES: normalized enrichment score.

ses. The measurement data were presented as the means ± standard error. Comparisons among the two groups were made using Student’s t-tests. A P-value < 0.05 was considered statistically significant.

Results

Identification of DEGs in ACP

The analysis of the GSE68015 dataset identified 75 DEGs (adjusted P < 0.001, logFC ≥ 7 or ≤ -7), comprising 64 upregulated genes and 11 downregulated genes in the ACP samples, compared with their expressions in healthy brain tissues.

GO and KEGG pathway enrichment analysis of DEGs in ACP using GSEA

GSEA was used to analyze GO and KEGG to explore the functions and pathways of the DEGs. The GO enrichment analysis showed that 3211/4136 gene sets were upregulated in ACP, 950 gene sets were significantly enriched at a nominal P-value < 0.05, and 334 gene sets were significantly enriched at a nominal P-value < 0.01. In addition, 925/4136 gene sets were downregulated in ACP, 125 gene sets were significantly enriched at a nominal P-value < 5%, and 30 gene sets were significantly enriched at a nominal P-value < 0.01. The most significant enrichments for both up- and down-regulated gene sets, in order of significance of the normalized enrichment score (NES), are listed in Table 1. Six significant enrichment plots are shown in Figure 1, such as “GO_REGULATION_OF_TRANSCRIPTION_ELONGATION_FROM_RNA_POLYMERASE_II_PROMOTER”, “GO_INTERMEDIATE_FILAMENT_BASED_PROCESS”, “GO_REGULATION_OF_ACTIVIN_RECEPTOR_SIGNALING_PATHWAY”, “GO_GLUTAMATE_RECEPTOR_BINDING”, “GO_TRANSPORT_VESICLE_MEMBRANE”, “GO_REGULATION_OF_RESPONSE_TO_FOOD”. Our GO enrichment analysis found that the gene sets upregulated in ACP were mainly associated with intermediate filaments and DNA binding, and the downregulated gene sets were frequently correlated with transport-vesicle membranes. The KEGG enrichment analysis showed that 130/168 gene sets were upregulated in ACP compared with their expressions in the healthy brain samples, 34 gene sets were significantly enriched at a nominal P-value < 5%, and nine gene sets were significantly enriched at a nominal P-value < 1%. There were 38/168 gene sets downregulated in ACP, and one gene set was significantly enriched at a nominal P-value < 5%. The top ten gene sets correlated with ACP, according to their NES, are shown in Table 2. Six significant enrichment plots are shown in Figure 2, such as “KEGG_PATHWAYS_IN_CANCER”, “KEGG_PATHWAYS_IN_CANCER”, “KEGG_BLADDER_CANCER”, “KEGG_CHRONIC_MYEOID_LEUKEMIA”, “KEGG_SMALL_CELL_LUNG_CANCER”, “KEGG_PATHOGENIC_ESCHERICHIA_COLI_INFECTION”.

The most significant enrichment plots: GO_INTERMEDIATE_FILAMENT_BASED_PROCESS. C. GO_REGULATION_OF_ACTIVIN_RECEPTOR_SIGNALING_PATHWAY. D. GO_GLUTAMATE_RECEPTOR_BINDING. E. GO_REGULATION_OF_RESPONSE_TO_FOOD. F. GO_VESICLE_DOCKING_INVOLVED_IN_EXOCYTOSIS.
Figure 2. Six significant enrichment plots from the pathway enrichment analyses of the DEGs in ACP, using GSEA. A. KEGG_PATHWAYS_IN_CANCER. B. KEGG_BLADDER_CANCER. C. KEGG_CHRONIC_MYELOID_LEUKEMIA. D. KEGG_PAXIMAL_TUBULE_BICARBONATE_Reclamat_i_n. E. KEGG_CARDIAC_MUSCLE_CONTRACTION. F. KEGG_ALANINE ASPARTATE_AND_GLUTAMATE_META_BoLISM.
Analysis of adamantinomatous craniopharyngioma

PROXIMAL_TUBULE_BICARBONATE_RECLAMATION. E. KEGG_CARDIAC_MUSCLE_CONTRACTION. F. KEGG_ALANINE ASPARTATE_AND_GLUTAMATE_METABOLISM.

BLADDER_CANCER", “KEGG_CHRONIC_MYELOID_LEUKEMIA”, “KEGG_PROXIMAL_TUBULE_BICARBONATE_RECLAMATION”, “KEGG_ALANINE ASPARTATE_AND_GLUTAMATE_METABOLISM”.

According to the KEGG pathway enrichment analysis, the gene sets upregulated in ACP were mostly associated with pathways linked with cancer, and the downregulated gene sets participated in the alanine, aspartate, and glutamate metabolism and neuroactive ligand receptor interactions.

PPI network construction and module analysis

The PPI network of the DEGs comprised 55 nodes and 162 edges, with a PPI enrichment P-value < 1.0e-16 (Figure 3). The cytoHubba plug-in for Cytoscape was used to screen the hub genes in the PPI network. A significant module, including 10 nodes and 45 edges, was generated from the PPI network of DEGs by MCODE (Figure 4). Five genes were identified as hub genes: CDH1, KRT19, KRT16, KRT5, and KRT14 (Figure 5). These five hub genes were screened out in the heatmap (Figure 6).

Discussion

There are different growth patterns of ACP, and within the sellar region, where this tumor occurs, there are many important structures, including the pituitary stalk, the optic nerve, the optic chiasm, and the hypothalamus [28-30]. Therefore, it remains a major challenge for neu-
Analysis of adamantinomatous craniopharyngioma

For this study, a total of 75 DEGs were identified, comprising 64 upregulated genes and 11 downregulated genes in the ACP samples compared with their expressions in the healthy brain tissues. The GSEA of DEGs provided a comprehensive overview of some of the major pathophysiological mechanisms involved in ACP, including intermediate filaments, transport vesicles, and membranes.

Intermediate filaments (IF) are one of the main types of protein filament that form the cytoskeleton [32, 33]. Many diseases are attributable to changes in cytoskeleton proteins, especially changes in the IF protein [34-36]. Gene deletions and the irregular expression of IF proteins can give rise to some types of tumors. Mu-Min Shao [37] concluded that cytokeratin (CK) immunohistochemistry plays an important role in the development of breast carcinoma. High CK7, CK8, CK18, and CK19 expression rates have also been observed in breast carcinoma. Furthermore, CK8 and CK20 play a very significant role in colorectal cancer, with high diagnostic and prognostic relevance [38]. In addition, Minsun Jung found that the CK20-positive state was associated with poorer prognoses and shorter survival times in patients with urinary bladder transitional cell carcinoma [39, 40]. The present study found that the expression of IF was noticeably upregulated in the patients with ACP compared with its expression in the healthy control group. Therefore, we infer that the over-expression of IF proteins might lead to ACP, which indicates IF could be a potential target for the treatment and prevention of ACP.

In this work we also built a PPI network and identified a total of five key genes: CDH1, KRT19, KRT16, KRT5, and KRT14. It was also...
discovered that the expression of CDH1 was enhanced in patients with ACP. CDH1 is the gene that codes for E-cadherin, which plays a role in suppressing tumor progression [41]. E-cadherin is a vital adhesion molecule. Tumor metastasis begins with a reduction in cell adhesion, which relies on adhesion molecules [42-44]. Jun Yang found that the expression of CDH1 is significantly upregulated in patients with ACP [20]. Other researchers have observed CDH1 mutations in patients with hereditary diffuse gastric cancer, including CDH1 gene deletions, E-cadherin deletions, and decreased E-cadherin expressions [45]. Hereditary diffuse gastric cancer is an autosomal dominant inherited disease associated with CDH1 gene mutations, and after lobular breast cancer is the second most frequent type of neoplasia [46]. CDH1 mutations, including gene deletions and decreases in expression, are also present in other frequently occurring tumors, such as ovarian tumors and colorectal tumors [47, 48].

However, our study has some limitations. First, the results are based only on a bioinformatics analysis, so laboratory studies will be necessary to verify the functions of the key genes, using numerous tissue samples from patients with ACP.

In conclusion, using a bioinformatics analysis, we identified 75 DEGs and five hub genes when we compared the ACP and the healthy brain samples. These genes may be involved in the occurrence and development of ACP, especially the genes involved with the intermediate filaments.

Acknowledgements

We are thankful to Qing-qing Wang (Tianjin Medical University) for her statistical assistance and suggestions. This work was funded by Capital Characteristic Clinical Application Research (grant number: Z161100000516-236).

Disclosure of conflict of interest

None.

Address correspondence to: Xin Yu, Department of Neurosurgery, Chinese PLA General Hospital-Sixth Medical Center, No. 6, Fucheng Road, Haidian District, Beijing 100037, P. R. China. Tel: +86-15210062313; Fax: 010-66958114; E-mail: yuxin37@sina.cn

References


Analysis of adamantinomatous craniopharyngioma


Analysis of adamantinomatous craniopharyngioma


[43] Reardon SN, King ML, MacLean JA 2nd, Mann JL, DeMayo FJ, Lydon JP and Hayashi K. CDH1 is essential for endometrial differentiation,


