Integrated bioinformatic analysis unveils significant genes and pathways in the pathogenesis of supratentorial primitive neuroectodermal tumor

Guang-Yu Wang1,*
Ling Li2,*
Bo Liu1
Xiao Han1
Chun-Hua Wang1
Ji-Wen Wang3

1Department of Neurosurgery, 2Department of Pediatrics, Qilu Children’s Hospital of Shandong University, Jinan, Shandong, 3Department of Neurology, Shanghai Jiaotong University School of Medicine, Pudong New District, Shanghai, People’s Republic of China

*These authors contributed equally to this work

Purpose: This study aimed to explore significant genes and pathways involved in the pathogenesis of supratentorial primitive neuroectodermal tumor (sPNET).

Materials and methods: Gene expression profile of GSE14295 was downloaded from publicly available Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) were screened out in primary sPNET samples compared with normal fetal and adult brain reference samples (sPNET vs fetal brain and sPNET vs adult brain). Pathway enrichment analysis of these DEGs was conducted, followed by protein–protein interaction (PPI) network construction and significant module selection. Additionally, transcription factors (TFs) regulating the common DEGs in the two comparison groups were identified, and the regulatory network was constructed.

Results: In total, 526 DEGs (99 up- and 427 downregulated) in sPNET vs fetal brain and 815 DEGs (200 up- and 615 downregulated) in sPNET vs adult brain were identified. DEGs in sPNET vs fetal brain and sPNET vs adult brain were associated with calcium signaling pathway, cell cycle, and p53 signaling pathway. CDK1, CDC20, BUB1B, and BUB1 were hub nodes in the PPI networks of DEGs in sPNET vs fetal brain and sPNET vs adult brain. Significant modules were extracted from the PPI networks. In addition, 64 upregulated and 200 downregulated overlapping DEGs were identified in both sPNET vs fetal brain and sPNET vs adult brain. The genes involved in the regulatory network upon overlapping DEGs and the TFs were correlated with calcium signaling pathway.

Conclusion: Calcium signaling pathway and several genes (CDK1, CDC20, BUB1B, and BUB1) may play important roles in the pathogenesis of sPNET.

Keywords: primitive neuroectodermal tumor, microarray analysis, protein-protein interaction, transcription factors

Introduction

Primitive neuroectodermal tumors (PNETs), which are primarily composed of undifferentiated or poorly differentiated round neuroepithelial cells, are highly malignant embryonal neoplasms arising mostly in children (although they can also occur in adults). PNETs include supratentorial primitive neuroectodermal tumor (sPNET; also called central nervous system [CNS]-PNET) and medulloblastoma. Although sPNET is histologically indistinguishable from medulloblastoma, there is growing evidence showing different molecular characteristics between sPNET and medulloblastoma.

The progress on sPNETs has been relatively modest compared with medulloblastoma, partly because of the scarcity of good material for biological studies. sPNETs are associated with poor prognosis, with a 5-year disease free survival of 15%–50%
depending on treatment modality. Additionally, up to 80% of the children still die from recurrent disease using current standard therapy. Therefore, gaining more insights into the underlying molecular mechanisms of sPNETs is of great importance.

In recent years, knowledge about the molecular events contributing to the development and progression of PNETs has increased dramatically. A study showed that SMARCB1/INI1 could be a potential tumor suppressor gene in CNS-PNET, with the loss of INI1 protein expression being a marker of poor prognosis. Koch et al reported the presence of catenin beta-1 (CTNNB1) mutations in a small subset of sporadic cerebellar and cerebral PNETs and suggested that WNT/wingless signaling pathway may be involved in the molecular pathogenesis of PNETs. In addition, the investigation of Phi et al showed that the proportion of phosphorylated signal transducer and activator of transcription 3α (STAT3α) relative to STAT3α in sPNET was significantly greater than that in medulloblastoma, which suggested that the JAK/STAT3 pathway was activated in sPNET. Besides, Rogers et al. showed that WNT/β-catenin pathway could be used as a potential prognostic marker of sPNET, which could be used to stratify disease risk for patients with sPNET. However, the exact molecular mechanisms responsible for pathogenesis of PNET, especially of sPNET, are still poorly understood.

In the present study, we downloaded a gene expression profile from publicly available database and carried out systematic bioinformatics analysis. Differentially expressed genes (DEGs) were screened out in primary sPNET samples compared with normal fetal and adult brain reference samples, respectively, forming two comparison groups (sPNET vs fetal brain and sPNET vs adult brain). Pathway enrichment analysis of these DEGs was conducted, followed by protein–protein interaction (PPI) network construction and significant module selection. Besides, transcription factors (TFs) regulating the common DEGs in the two comparison groups were identified, and the regulatory network was constructed. The aim of this study was to identify significant genes and pathways involved in the pathogenesis of sPNET, which may facilitate the understanding of molecular mechanisms underlying sPNET development and open up new avenues for developing novel targets for diagnostic and therapeutic intervention.

Materials and methods

Microarray data

Gene expression profile of GSE14295 (accession number), deposited by Li et al., was downloaded from the public repository, Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/). The platform is GPL6102, Illumina human-6 v2.0 expression beadchip. This dataset consisted of 33 primary sPNET samples with good RNA quality. Additionally, seven normal fetal and seven adult brain reference samples were also included in GSE14295.

Data preprocessing and screening of DEGs

The normalized gene expression matrix and platform annotation profiles were downloaded. Based on the annotation information, the probes were matched to the gene symbols. For multiple probes corresponding to one gene symbol, we calculated the gene expression values by averaging across all the matched probe sets of a gene. Finally, 20,057 gene expression matrices were obtained.

After data preprocessing, we used the bioconductor limma package to calculate the level of gene differential expression in the two groups, sPNET vs fetal brain and sPNET vs adult brain. Benjamini–Hochberg (BH) multiple testing correction was applied to control the false-positive results, adjusting the raw p-values into false discovery rate (FDR). A combination of FDR <0.05 and the log2 fold change (FC) ≥2 was used as the stringent cutoff to determine the significance of gene expression difference. Additionally, hierarchical clustering of the identified DEGs was carried out to construct a dendrogram based on their expression values.

Pathway enrichment analysis of DEGs

The Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics resources (http://david.niaid.nih.gov) can enable the researchers to interpret the biological mechanisms associated with large gene lists and is the most common tool for functional enrichment analysis. In this study, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched by DEGs in sPNET vs fetal brain and sPNET vs adult brain were identified using DAVID (version 6.8). KEGG pathways with DEGs ≥2 and p-value <0.05 were considered to be significantly enriched.

PPI network construction of DEGs and module selection

The Search Tool for the Retrieval of Interacting Genes (STRING; http://www.string-db.org/) provides a critical assessment and integration of PPIs, including known and predicted associations. Using STRING, users can inspect the interaction evidence by adjusting interaction confidence scores. In this study, STRING was applied to select the interactions among the identified DEGs in sPNET vs fetal brain.
brain and sPNET vs adult brain, respectively. Besides, the confidence score >0.9 (indicating highest confidence) was chosen for the PPI network construction. Besides, Cytoscape software was used to create network visualizations. Analysis of topological properties of nodes in the network helps to describe the importance of nodes in the PPI network. Three methods including degree centrality, betweenness centrality, and closeness centrality were used to analyze network topological features of nodes in the PPI network using the CytoNCA plug-in (http://apps.cytoscape.org/apps/cytonca) in Cytoscape software. Hub nodes with higher topology scores were identified. Significant modules (densely connected subnetwork regions) in the PPI network were selected using the Cytoscape plug-in MCODE. Modules were selected based on the MCODE cluster score (i.e., connectivity density). In this study, we selected the modules with a score >6.

Quantitative Venn diagram distribution analysis of DEGs in two comparison groups

VennPlex (http://www.ipr.nia.nih.gov/bioinformatics/vennpplex.html), with its novel feature of visualizing contraregulated data points, can be used for comparing and visualizing datasets with differentially regulated data points, facilitating a more comprehensive and interconnected appreciation. In this study, Venn diagrams were constructed to identify common and contraregulated genes in PNET vs fetal brain and PNET vs adult brain using VennPlex.

Identification of TFs potentially regulating the common DEGs and construction of regulatory network

iRegulon (http://iregulon.aertslab.org), a Cytoscape plug-in, provides position weight matrices (PWMs) that have been collected from various sources and species and enables to construct regulatory network by linking them to candidate TFs. In this study, TFs potentially regulating the common DEGs identified in sPNET vs fetal brain and sPNET vs adult brain were examined using iRegulon. The lists of common DEGs identified were subjected to iRegulon and used to predict the TFs with the following parameters – minimum identity between orthologous genes: 0.05 and maximum FDR on motif similarity: 0.001. The predicted regulatory relationship with normalized enrichment scores (NEs) >3 was used for further analysis. The corresponding regulatory network was constructed. Moreover, genes enriched in the regulatory network were subjected to pathway enrichment analysis.

Results

Analysis of DEGs

After preprocessing, a stringent differential gene expression analysis (FDR <0.05 and |log2 FC| ≥2) revealed 526 DEGs in sPNET vs fetal brain, of which 99 genes displayed increased expression and 427 genes displayed decreased expression. In addition, we got 815 DEGs in sPNET samples compared with adult brain samples, which included 200 upregulated genes and 615 downregulated genes. The expression values of these DEGs were hierarchical as shown in Figure 1A and B.

Pathway enrichment analysis of DEGs

Pathway annotation was used to screen out the altered biological functions arising from the identified DEGs. Pathway enrichment of the up- and downregulated genes was carried out separately. As shown in Figure 1C and D, the downregulated genes in sPNET vs adult brain were mainly enriched in 15 pathways, including calcium signaling pathway and tight junction. The upregulated genes in sPNET vs adult brain were mainly enriched in seven pathways, such as cell cycle and p53 signaling pathway. By contrast, the downregulated genes in sPNET vs fetal brain were mainly enriched in 10 pathways, which included neuroactive ligand–receptor interaction and the calcium signaling pathway. The upregulated genes in sPNET vs adult brain were associated with 24 pathways, such as cell cycle and p53 signaling pathway.

PPI network construction and module selection

The PPI networks upon the DEGs in sPNET vs fetal brain and sPNET vs adult brain are shown in Figures 2 and 3, respectively. The PPI network of DEGs in PNET vs fetal brain was composed of 155 nodes (proteins) and 241 edges (interactions; Figure 2). The PPI network of DEGs in PNET vs adult brain consisted of 290 nodes and 733 edges (Figure 3). Top 20 nodes with higher topology scores in the PPI networks are shown in Table 1. The results of topological analysis showed that cyclin-dependent kinase 1 (CDK1), cell division cycle 20 (CDC20), adenylate cyclase 2 (ADCY2), adenylate cyclase 8 (ADCY8), MAD2 mitotic arrest deficient-like 1 (MAD2L1), aurora kinase A (AURKA), BUB1 mitotic checkpoint serine/threonine kinase B (BUB1B), and BUB1 mitotic checkpoint serine/threonine kinase B (BUB1B) were hub proteins with a higher degree centrality, betweenness centrality, and closeness centrality in the PPI network of DEGs in PNET vs fetal brain. On the other hand, CDK1, aurora kinase B (AURKB), CDC20, BUB1B, and BUB1 were hub proteins with a higher degree centrality, betweenness centrality and closeness centrality in the PPI network of DEGs in PNET vs adult brain.
Figure 1 Hierarchical clustering and pathway enrichment analysis of differentially expressed genes (DEGs) in primitive neuroectodermal tumor (PNET) vs fetal brain and PNET vs adult brain.

Notes: (A) Heat-map overview of DEGs in PNET vs fetal brain. Red represents upregulated genes, and green indicates downregulated genes. (B) Heat-map overview of DEGs in PNET vs adult brain. Red represents upregulated genes, and green indicates downregulated genes. (C) Pathway enrichment analysis of DEGs in PNET vs fetal brain. Up represents upregulated genes, and down indicates downregulated genes. (D) Pathway enrichment analysis of DEGs in PNET vs adult brain. Up represents upregulated genes, and down indicates downregulated genes.

In addition, two significant modules were selected from the PPI network of DEGs in sPNET vs fetal brain (Figure 4A and B), and two significant modules were selected from the PPI network of DEGs in sPNET vs adult brain (Figure 4C and D). Besides, we performed pathway enrichment analysis of genes enriched in the significant modules (Table 2; A, B, C, and D corresponded to the modules in Figure 4). The results showed that the upregulated genes in sPNET vs fetal brain, such as CDC20 and MAD2LI1, were associated with cell cycle. The downregulated genes in PNET vs fetal brain, such as metabotropic glutamate receptor 3 (GRM3), ADCY2, and ADCY8, were involved in neuroactive ligand–receptor interaction and glutamatergic synapse. By contrast, the downregulated genes in sPNET vs adult brain such as ACTN2 and MYH6 were concerned with tight junction. The upregulated genes in PNET vs adult brain such as CDC20, BUB1, and BUB1B were correlated with cell cycle.
Figure 2 Protein–protein interaction (PPI) network of differentially expressed genes (DEGs) in primitive neuroectodermal tumor (PNET) vs fetal brain.

Note: White nodes stand for the upregulated DEGs, while gray nodes stand for the downregulated genes.

Figure 3 Protein–protein interaction (PPI) network of differentially expressed genes (DEGs) in primitive neuroectodermal tumor (PNET) vs adult brain.

Note: White nodes stand for the upregulated DEGs, while gray nodes stand for the downregulated genes.
Table 1 Top 20 genes that are evaluated by degree centrality, betweenness centrality, and closeness centrality in the protein–protein interaction (PPI) networks of differentially expressed genes (DEGs) in primitive neuroectodermal tumor (PNET) vs fetal brain and PNET vs adult brain

<table>
<thead>
<tr>
<th>PNET_fetal PPI</th>
<th>PNET_adult PPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node</td>
<td>Degree</td>
</tr>
<tr>
<td>CDK1</td>
<td>27</td>
</tr>
<tr>
<td>CDC20</td>
<td>18</td>
</tr>
<tr>
<td>ACDY2</td>
<td>17</td>
</tr>
<tr>
<td>ACDY1</td>
<td>15</td>
</tr>
<tr>
<td>MAD2L1</td>
<td>14</td>
</tr>
<tr>
<td>AURKA</td>
<td>14</td>
</tr>
<tr>
<td>BUB1B</td>
<td>13</td>
</tr>
<tr>
<td>NPY</td>
<td>13</td>
</tr>
<tr>
<td>BUB1</td>
<td>12</td>
</tr>
<tr>
<td>SST</td>
<td>12</td>
</tr>
<tr>
<td>SSTR1</td>
<td>12</td>
</tr>
<tr>
<td>ADY1</td>
<td>12</td>
</tr>
<tr>
<td>CENPE</td>
<td>11</td>
</tr>
<tr>
<td>NDC80</td>
<td>11</td>
</tr>
<tr>
<td>GRM3</td>
<td>11</td>
</tr>
<tr>
<td>CNR1</td>
<td>11</td>
</tr>
<tr>
<td>GRM2</td>
<td>11</td>
</tr>
<tr>
<td>CAMK2B</td>
<td>10</td>
</tr>
<tr>
<td>CENPN</td>
<td>10</td>
</tr>
<tr>
<td>CENPF</td>
<td>9</td>
</tr>
</tbody>
</table>

Identification of common DEGs in PNET vs fetal brain and PNET vs adult brain

As shown in Figure 4E, 64 common upregulated genes and 200 common downregulated genes were identified in sPNET vs fetal brain and PNET vs adult brain. Besides, one gene was identified to be a contraregulated gene. These common DEGs identified were all with |log2 FC| >2.

Identification of TFs potentially regulating the common DEGs and construction of regulatory network

Based on the common DEGs identified using VennPlex, iRegulon identified three TFs potentially regulating these common DEGs in sPNET vs fetal brain and sPNET vs adult brain. The regulatory network was constructed as shown in Figure 5A, consisting of 114 common DEGs, three TFs (transcription factor 12 [TCF12], zinc finger protein 260 [ZNF260], and glycerol-3-phosphate dehydrogenase 1 [GPD1]), and 175 interaction pairs. The |log2 FC| values of these three TFs (TCF12, ZNF260, and GPD1) in sPNET vs adult brain were all >2. The log2 FC values of TCF12, ZNF260, and GPD1 in sPNET vs fetal brain were 1.9880, 1.2513, and −0.6292, respectively. On the other hand, as shown in Figure 5B, the genes enriched in the regulatory network were involved in 28 pathways, such as hsa04024: cyclic adenosine monophosphate (cAMP) signaling pathway and hsa04020: calcium signaling pathway.

Discussion

To advance sPNET therapeutics, it is of significance to delineate the cellular and molecular pathogenesis of sPNET. In the current study, we identified 526 DEGs (99 up- and 427 downregulated) in sPNET vs fetal brain and 815 DEGs (200 up- and 615 downregulated) in sPNET vs adult brain. DEGs in sPNET vs fetal brain and sPNET vs adult brain were associated with calcium signaling pathway, cell cycle, and p53 signaling pathway. CDK1, CDC20, BUB1B, and BUB1 were hub proteins in the PPI network of DEGs in sPNET vs fetal brain and sPNET vs adult brain. Two modules enriched by DEGs in sPNET vs fetal brain and two modules enriched by DEGs in sPNET vs adult brain were extracted from the PPI networks. A total of 264 overlapping DEGs (64 upregulated and 200 downregulated) were identified in both sPNET vs fetal brain and sPNET vs adult brain. The genes involved in the regulatory network, which was constructed based on the overlapping DEGs and the identified TFs, were significantly correlated with the calcium signaling pathway and cAMP signaling pathway.

Ca2+ signals impact nearly every aspect of cellular life and are a core regulator of cellular responses to the environment. Evidence have shown that altered expression of specific Ca2+
Significant genes and pathways in sPNET

pumps and channels are characterizing features of some diseases and cancers. The findings of Mattson and Chan demonstrated that perturbed cellular calcium homeostasis played a pivotal role in the pathogenesis of Alzheimer’s disease. Besides, several known molecular players such as stromal interaction molecule 1 (STIM1), calcium release-activated calcium channel protein 1 (ORAI1), and transient receptor potential (TRP) channels, which are involved in cellular Ca^{2+} homeostasis, have been implicated in tumor cell migration and tumor metastasis. Davis et al reported the role of the intracellular calcium signal in the induction of epithelial–mesenchymal transition (EMT) in human breast cancer cells. Moreover, specific Ca^{2+} signals decoded by different Ca^{2+} sensor proteins allow a fine-tuning of the Ca^{2+} signaling, and Ca^{2+} signaling pathway is involved in early neural development. In this study, the overlapping DEGs of sPNET vs fetal brain and sPNET vs adult brain were significantly concerned with calcium signaling pathway,
suggesting abnormalities in calcium regulation. In this context, it is surmised that calcium signaling pathway may play a prominent role in pathology of sPNET and may shed light on the discovery of new therapeutic targets of sPNET.

The cAMP signaling pathway is an important intracellular system for signal transmission that is involved in many cell types. Numerous studies indicate cAMP as central players underlying the pathogenesis of many diseases and show treatment strategies targeting cAMP signaling. A study demonstrated that a cAMP-dependent pathway was activated during astrocyte differentiation. Moreover, activation of cAMP signaling pathway is a key determinant of neuronal differentiation and plasticity. In this study, we found that overlapping DEGs in sPNET vs fetal brain and sPNET vs adult brain were involved in the cAMP signaling pathway. In this context, we suggested alteration of the cAMP signaling pathway as a potential mechanism in sPNET pathogenesis.

The protein encoded by CDK1 is a member of the Ser/Thr protein kinase family and is a central regulator that drives cells through the G2 phase and mitosis. The cyclin-dependent kinases, such as cyclin-dependent kinase 5 (Cdk5), is shown to play important roles in brain development and disease. A recent published study demonstrated that CDK1 could induce the activation of pre-mitotic nuclear envelope dynein recruitment and apical nuclear migration in neural stem cells. Besides, multiple CDK/CYCLIND genes are identified to be amplified in sPNET and medulloblastoma. In addition, CDK1 was found to be one of hub proteins in the PPI network of DEGs in sPNET vs fetal brain and sPNET vs adult brain. Thus, we suggested that dysregulation of CDK1 expression may be correlated with the development of sPNET through participating in cell-cycle regulation. However, more related investigations are needed to check the result.

On the other hand, CDC20 was another hub protein of the PPI networks in this study. CDC20 is an essential cell-cycle regulator interacting with several other proteins at multiple points in the cell cycle. Ras association domain family member 1 (RASSF1A) can inhibit APC-CDC20 activity through its D-box domains at early mitosis. Evidence demonstrated that the inactivation of the RASSF1A gene may play an essential role in the tumorigenesis of PNET of the central nervous system. Besides, CDC20 was identified to be upregulated in samples of sPNET and was a hub protein with a higher degree centrality, betweenness centrality, and closeness centrality in the PPI networks. Taken together, these data provided a potential link between CDC20 and the pathogenesis of sPNET.

Moreover, BUB1B and BUB1 were also hub proteins with higher topology scores in the PPI network of DEGs in sPNET vs fetal brain and sPNET vs adult brain. BUB1B and BUB1 are essential components of the mitotic checkpoint. The human brain tumor, glioblastoma multiforme, was reported for the requirement for BUB1 to suppress lethal consequences of altered kinetochore function. Morales et al showed that inhibition of BUB1 and BUBR1 decreased cell proliferation and colony formation in pediatric glioblastoma cells and enhanced radiation sensitivity. In this study, BUB1B and BUB1 were upregulated DEGs. It is tempting...
to speculate that BUB1B and BUB1 play critical roles in the development and progression of sPNET.

Despite merits, our study had several limitations. First, the sample size was small, which may limit our power to detect real differences in gene expression. Further studies with a larger simple size should be performed. Second, the study lacked experimental validations and cross-checking with other datasets. Pooling data from similar studies and

Figure 5 Regulatory network construction and pathway enrichment analysis of genes in the regulatory network.
Notes: (A) A regulatory network based on common DEGs and three transcription factors (TFs). Elliptical nodes indicate the common DEGs in sPNET vs fetal brain and sPNET vs adult brain. Triangle represents TF. (B) Pathways enriched by the genes in the regulatory.
performing experimental validations will be the focus of our future investigations.

Conclusion
Our study suggested that calcium signaling pathway, cAMP signaling pathway, and several DEGs (CDK1, CDC20, BUB1B, and BUB1) may play important roles in the pathogenesis of sPNET. Further research is warranted to determine these findings.

Acknowledgments
This work was supported by the Special Foundation for Taishan Scholars (No ts20110814) and Science and Technology Program (2014GSF118179), a project of Shandong province.

Disclosure
The authors report no conflicts of interest in this work.

References