**Title:** A differential pathway network method identifies pivotal pathways associated with pediatric pilocytic astrocytoma

**Running title:** Pivotal pathways in pediatric PA

Authors: Qiao-Ling Wu1, Jun-Bo Yang2, Hai-Xia Liu3, Xia Lin4, Xiu-Fang Fan1, Zhen-Feng Liu5\*

1Department of Pediatrics, Jinan Maternity and Child Care Hospital Affiliated to Shandong University, Jinan 250001, China.

2Department of Pediatrics, The 7th People's Hospital of Jinan, Jinan 251400, China.

3Department of Pathology, Jinan Maternity and Child Care Hospital Affiliated to Shandong University, Jinan 250001, China.

4Pediatric Intensive Care Unit, Qilu Children’s Hospital of Shandong University, Jinan 250022, China.

5Department of Traditional Chinese Medicine, Jinan Maternity and Child Care Hospital Affiliated to Shandong University, Jinan 250001, China.

**\* Corresponding author:** Zhen-Feng Liu; E-mail: sdzhenfeng@126.com; Tel: +86-0531-89029457; Fax: +86-0531-82060983; Cellphone: +86-13853169280; Mailing address: Jinan Maternity and Child Care Hospital Affiliated to Shandong University, No. 2 Jianguo-Xiaojingsan Road, Jinan 250001, China.

**Abstract**

This paper aims to identify the underlying pathways associated with the development and progression of pediatric pilocytic astrocytomas (PA) utilizing a differential pathway network (DPN) strategy. Thecurrent study proposed a novel form of molecular network, i.e., DPN, where the protein-protein interaction (PPI) network and the pathway crosstalks were integrated with the differential expression of pathways. By data preprocessing, a total of 1505 pathways with the genes size of more than 5 and less than 100 were selected, of which 1972 pathway crosstalks with weight value > 1.2 were identified as differential pathway interactions. The pathways and the corresponding crosstalks were visualized via Cytoscape to construct the final DPN, covering 956 pathways and 1972 pathway interactions. Finally, we identified 10 pivotal pathways based on DPN topological analysis. In this study, we propose a novel differential pathway network that covers disease-related differential pathway interactions. From the DPN, we successfully identified several pivotal pathways that might be associated with the development and progression of pediatric of PA.

**Key words:** Pediatricpilocytic astrocytomas; differential pathway network; protein-protein interaction; topological analysis

**Introduction**

Pilocytic astrocytomas (PA) are known to be the major brain solid neoplasms in children and teenagers.([1989](#_ENREF_1); [Scheithauer et al., 2007](#_ENREF_21); [Ward et al., 2014](#_ENREF_24)) They usually arise in the cerebellum, hypothalamic region or the optic chiasm, but may occur in any area where astrocytes are present, including the cerebral hemispheres and the spinal cord. These tumors are usually slowly grown and benign.([Huang et al., 2005](#_ENREF_11)) Despite the overall good prognosis of PA, up to 20% of patients will have a poor outcome, with recurrence, growth of incompletely resected lesions, or dissemination through the cerebrospinal fluid, and ultimately death due to disease.([Scheithauer et al., 2007](#_ENREF_21); [Jones et al., 2012](#_ENREF_14)) Because of the age of people diagnosed with PAs, the treating medical team will often try to avoid radiotherapy and chemotherapy in order to avoid damage to the developing brain. Meanwhile, the tumor location could prohibit access to the neoplasm and lead to incomplete or no resection at all. Equipping with better knowledge on biological markers and pathology of pediatric PA may provide new possible targeted treatments.

Rapid advances on high-throughput technologies have brought unprecedented opportunities to measure molecular behaviors underlying the phenotype changes at a genome-wide level and to transform the data into a meaningful biological phenomenon.([Ono and Han, 2000](#_ENREF_20); [Logan and Nusse, 2004](#_ENREF_17); [Inoki et al., 2005](#_ENREF_12)) To date, a variety of methods have been proposed to identify molecular biomarkers for complex diseases. A crosstalk among pathways is mean of regulatory interaction and describes phenotype differences from the pathway interaction viewpoint.([Colaprico et al., 2015](#_ENREF_8)) Traditional methods mainly identify differential pathways that are dramatically different in disease state and normal state, and a natural strategy looking for a pathway crosstalk is to identify the common genes between two pathways.([Zhou et al., 2015](#_ENREF_26); [Zhuang et al., 2015a](#_ENREF_27)) While, these methods ignore the dysregulated interactions among pathways to model underlying disease development and progression. Pathways function in a cooperative way (as physical or biochemical interactions) but not in isolation, perturbed pathway interactions might cause final molecular dysregulation underlying complex diseases. Therefore, we investigate the differential pathway interactions that are significantly changed from wildtype to disease condition.([Bradley et al., 2008](#_ENREF_6))

Given the concept of differential pathway interaction, this paper aims to conduct a systematical analysis to identify the pivotal pathways in pediatric PA, by developing a novel form of molecular network, differential pathway network (DPN). To achieve this, the protein-protein interaction (PPI) data and pathway interactions were integrated with the differential expression of pathways in pediatric PA. Pathway crosstalks were subsequently randomly extracted based on gene-gene interactions, and a weight value that respected the differential expression was assigned to each crosstalk using the Spearman correlation coefficient (SCC). The pathways and the corresponding crosstalks then formed the final DPN. This paper would offer a novel way to predict markers in disease in an accurate manner. The details was listed in Methods section.

**Materials and Methods**

**Data**

*Gene expression data*

# To identify differential pathways in pediatric PA, we retrieved the microarray data of pediatric PA from ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/), under the accession number of E-GEOD-44971. This dataset was presented on A-AFFY-44 - Affymetrix GeneChip Human Genome U133 Plus2.0 platform. The gene expression data were gained from 49 PA tumor samples and 9 normal cerebellum samples.([Lambert et al., 2013](#_ENREF_16)) The microarray data and annotation files were downloaded for further analysis. The gene expression profile on probe level was converted into gene symbol level, as well as wiping off the duplicated symbols. Finally, we obtained a total of 20,109 genes in expression profile.

*PPI data*

All human PPI data were downloaded from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, http://string-db.org/) database which provides a comprehensive, yet quality-controlled collection of protein association data for a large number of organisms.([von Mering et al., 2005](#_ENREF_23)) There are a total of 1,048,576 human gene interactions, and a total of 787,896 interactions (covering 16,730 genes) were remained after removing self-loops and duplicated interactions. We selected the interactions whose two nodes were both existed in the expression profile for further analysis, covering 734,191 interactions and 15,246 genes.

*Pathway data*

# Information from gene sets representing all human biological pathways was downloaded from the Reactome database ([http://www.reactome.org](http://www.reactome.org/)) which is an open source curated bioinformatics database of human pathways and reactions.([Croft et al., 2011](#_ENREF_9)) Since pathways with too many genes might be too complex to understand, and pathways with too few genes may have not sufficient biological content,([Ahn et al., 2014](#_ENREF_3)) thus the pathways with gene size more than 5 and less than 100 genes were selected. In this study, 1505 of 1675 pathways were collected for further analysis.

**Differential pathway interactions and DNP**

In biological system, pathways function together in a highly coordinated way to respond appropriately to various stimuli. In order to evaluate the interactions between pathways, we first randomly constructed interactions between genes in the pathways based on PPI data and pathway data. If gene interactions could be detected between two pathways, these two pathways were inclined to interact with each other, that is to say, these two pathways had a crosstalk. For the pathway 1 and pathway 2, we screened all genes relationships between them, and defined the intersection of all genes relationships as the pathway interaction between pathway1 and pathway 2.

To detected the differential pathway interactions in pediatric PA relative to normal condition, a weight value was assigned to each crosstalk. In the present study, the weight value of a pathway crosstalk was defined as the absolute differences of SCC between pediatric PA and normal controls. To achieve this, we first calculated the absolute differences of SCC of each gene interaction between pediatric PA and normal controls. The mean value of absolute difference for all intersected gene interactions was denoted as the absolute differences of a pathway crosstalk, i.e. the weight value of the pathway crosstalk. Then, we analyzed the weight value of pathway interaction between pathway 1 and 3 according to the above methods. The rest interactions and weight values among the pathways may be deduced by analogy. After that, we obtained all weight values of pathway interactions. In this article, the interactions with weight values greater than 1.2 were considered as differential pathway interactions. Finally, DPN was visualized via Cytoscape software based on differential pathway crosstalks and their corresponding weight values.

# Topology analysis of DPN

# To further gain the pivotal pathways related to pediatric PA from the DPN, we conducted on the topology analysis for the DPN based on the nodes degree.([Magoni, 2002](#_ENREF_18)) The topology analysis mainly contains degree, closeness, betweenness and stress, in which degree quantifies the local topology of each node by summing up the number of its adjacent nodes in the network.([Zhuang et al., 2015b](#_ENREF_28)) Degree gives a simple count of the number of interactions of a given node and is particularly useful to identify key players in biological processes. In the present study, we picked the pathways with the top 1% degree as the pivotal pathway in pediatric PA.

**Results**

**Data acquisition**

After data preprocessing, we obtained a total of 20,109 genes, 734,191 PPI (covering 15,246 genes) and 1505 pathways from the expression profile, STRING and Reactome database, respectively. To evaluate the interactions between pathways, we selected all possible pathway crosstalks. These 1505 pathways could form 816,000 pathway interactions. For convenience, each pathway was given a corresponding ID number in this article (as shown in supplementary material **Table S1**). Hereafter, the pathways were represented by these ID number.

**Differential pathway interactions**

In the current study, the absolute difference of SCC was computed to investigate the strength of interactions among pathways, which was defined as the weight value of pathway interaction. There were 816,000 pathway interactions among 1505 pathways. Each interaction had a weight value. The larger the weight value, the more difference of pathway interaction between disease and normal conditions. The distribution of weight values of all interactions was calculated, as shown in **Figure 1.** Among 816,000 pathway interactions, a total of 1972 differential pathway interactions (covering 965 pathways) were identified under the threshold value of weight value greater than 1.2.

F:\成型文章\2017年12月\额外投稿文章\田慧芬 2\田慧芬\1214-尹娟\1214投稿\Genetics and Molecular Research\Figure 1.tif

**Figure 1** The frequency distribution of weight values of pathway interactions in pediatric pilocytic astrocytomas.

Based on these differential pathway interactions and their corresponding weight values, the DPN was constructed via Cytoscape software (**Figure 2**). Different from gene network, in DPN, the node was pathway and the edge was the interaction between two pathways. Each edge had a weight value that represented the difference value of pathway interaction in disease condition relative to normal condition.

F:\成型文章\2017年12月\额外投稿文章\田慧芬 2\田慧芬\1214-尹娟\1214投稿\Genetics and Molecular Research\Figure 2.tif

**Figure 2** The differential pathway network for pediatric pilocytic astrocytomas. Node refers to pathway, and edge represents the interaction between two pathways. Red nodes were pivotal pathways. In this network, each pathway is given an ID number (see **Table S1**).

**Pivotal pathway**

From the DPN, we performed a degree centrality analysis to gain the pivotal pathways associated with pediatric PA. By ranking node degrees in descending order, the pathways with the top 1% degree distribution were considered as the pivotal pathways in pediatric PA. In the present study, we identified 10 pivotal pathways from DPN (**Table 1**). The pivotal pathway in the network interacted with more other pathways, indicating a more important role of pivotal pathway in regulatory network. The most two pivotal pathways were synthesis of 12-eicosatetraenoic acid derivatives (ID: 1285, degree = 36), and synthesis of 5-eicosatetraenoic acids (ID: 1287, degree = 34).

**Table 1.** The pivotal pathways in pediatric pilocytic astrocytomas

|  |  |  |
| --- | --- | --- |
| ID | Pathways | Degree |
| 1285 | Synthesis of 12-eicosatetraenoic acid derivatives | 36 |
| 1287 | Synthesis of 5-eicosatetraenoic acids | 34 |
| 1376 | Toxicity of botulinum toxin type B (BoNT/B) | 31 |
| 1381 | Toxicity of botulinum toxin type G (BoNT/G) | 31 |
| 203 | Catecholamine biosynthesis | 30 |
| 426 | Establishment of Sister Chromatid Cohesion | 25 |
| 195 | Cam-PDE 1 activation | 24 |
| 1286 | Synthesis of 15-eicosatetraenoic acid derivatives | 23 |
| 251 | Cohesin Loading onto Chromatin | 22 |
| 1173 | Sema4D mediated inhibition of cell attachment and migration | 21 |

From the network, it is obvious that one pivotal pathway could regulate many other pathways, and one pathway also could be regulated by two or more pivotal pathways. Further analysis showed that the pivotal pathways were decentralized in the network, but not agminated in the central of DPN, and formed three clusters by regulating a series of common pathways. For example, the three pathways of synthesis of 12-eicosatetraenoic acid derivatives (ID: 1285), Synthesis of 15-eicosatetraenoic acid derivatives (ID: 1286) and synthesis of 5-eicosatetraenoic acids (ID: 1287) were collaborative closely by regulating a series of common pathways (**Figure 3**). Detailedly, total 21 common pathways were regulated by these three pivotal pathways, indicating their strong relationships among each other and their important roles in network regulation.

F:\成型文章\2017年12月\额外投稿文章\田慧芬 2\田慧芬\1214-尹娟\1214投稿\Genetics and Molecular Research\Figure 3.tif

**Figure 3** Sub-network of three pivotal pathways (ID: 1285-synthesis of 12-eicosatetraenoic acid derivatives, ID: 1287-synthesis of 5-eicosatetraenoic acids and ID: 1286-synthesis of 15-eicosatetraenoic acid derivatives) and their adjacent pathways. Node refers to pathway, and edge represents the interaction between two pathways. Red nodes were pivotal pathways. In this network, each pathway is given an ID number (see **Table S1**).

**Discussion**

In the present study, we developed a novel form of molecular network, DPN, to systemically characterize the phenotype differences of PA samples, and provided biological insights into the development and progression of PA at the pathway network level. PA is the most common primary malignant brain tumor in the central nervous system.([Wen and Kesari, 2008](#_ENREF_25)) Despite new biological insights and therapeutic advances, the general prognosis for PA patients remains poor. A better understanding of the molecular markers of the disease is very important for early diagnosis, appropriate treatment and improved prognosis of patients with astrocytoma. Here we developed the pathway network perturbation model and uncovered the differential expression pathway interactions.

Various evidence has demonstrated biomolecules involved in the pathogenesis of PA,([Addo-Yobo et al., 2006](#_ENREF_2); [Jones et al., 2008](#_ENREF_15); [Becker et al., 2015](#_ENREF_4)) however its underlying mechanism remains obscure. It might be simplistic to explain the disease based on the role of a few genes only or the pathways alone. Owing to the rapid development of the human interactome knowledge base, network-based approach has become increasingly powerful for the study of disease mechanism.([del Sol et al., 2010](#_ENREF_10)) A variety of calculation methods have been put forward to detecting the disease related network, such as co-expression network,([Miller et al., 2008](#_ENREF_19)) PPI network.([Tian et al., 2014](#_ENREF_22)) Network-level strategy is indispensable to address unknown mechanism of PA based on the disease-associated pathways. To our best knowledge, this is the first time to construct a differential pathway network based on the overlapping pathway interactions and identify the pivotal pathways in PA. Our approach inferred the network strategy by identifying differential pathway interactions. Using a [topology](javascript:void(0);) [analysis](javascript:void(0);) of pathway network, we identified 10 pivotal pathways with the highest degree scores.

From the pathway network, we noticed that three pivotal pathways and their adjacent pathways form a cluster. These three pivotal pathways were synthesis of 12-eicosatetraenoic acid derivatives (ID: 1285, degree = 36), synthesis of 5-eicosatetraenoic acids (ID: 1287, degree = 34) and synthesis of 15-eicosatetraenoic acid derivatives (ID: 1286, degree = 23). These three pivotal pathways were all related to the synthesis of eicosatetraenoic acid, and were collaborative closely by regulating 21 common pathways. Eicosatetraenoic acid designates any straight chain 20:4 [fatty acid](https://en.wikipedia.org/wiki/Fatty_acid), and appears to act as dual inhibitor of arachidonic acid oxygenation by cyclooxygenase and lipoxygenase pathway.([Bierer and Bui, 2002](#_ENREF_5)) Lipoxygenase is a key enzyme that might promote carcinogenesis. Inhibition of the lipoxygenase pathways clearly has chemopreventive effects on various cancers.([Chen et al., 2006](#_ENREF_7)) Moreover, previous study has demonstrated the suppression of astrocytomas growth by inhibitors of the lipoxygenase pathway.([Ishii et al., 2008](#_ENREF_13)) In this study, these three eicosatetraenoic acid-related pathways were not only pivotal pathways with the highest degree values, but also effected many other pathways in a collaborative way. We speculated that the synthesis of eicosatetraenoic acid and its derivatives might be associated with the development and progression of pediatric PA.

Our DPN scheme proposes a novel way to predict disease-related pathways and pathway interactions, which identifies pathway network biomarkers by integrating gene expression, PPI and pathway data. This DPN scheme increased the molecular knowledge of PA, which might be beneficial for the further development of accurate therapeutics.

**Acknowledgement**There is no fund to support this study.  
**Conflict of interest**No conflict of interest to declare

**Reference**

Listed N (1989) Long term trends in cancer mortality rates from 1955 to 1987 in Japan. The Bureau of Vital Statistics, Ministry of Health and Welfare. Jpn J Clin Oncol 19: 305-317.

Addo-Yobo SO, Straessle J, Anwar A, Donson AM, Kleinschmidt-Demasters B and Foreman NK (2006) Paired overexpression of ErbB3 and Sox10 in pilocytic astrocytoma. Journal of Neuropathology & Experimental Neurology 65: 769-775.

Ahn T, Lee E, Huh N and Park T (2014) Personalized identification of altered pathways in cancer using accumulated normal tissue data. Bioinformatics 30: i422-i429.

Becker AP, Scapulatempo-Neto C, Menezes WP, Clara C, Machado HR, Oliveira RS, Neder L and Reis RM (2015) Expression of Methylthioadenosine Phosphorylase (MTAP) in Pilocytic Astrocytomas. Pathobiology 82: 84-89.

Bierer TL and Bui LM (2002) Improvement of arthritic signs in dogs fed green-lipped mussel (Perna canaliculus). J Nutr 132: 1634S-1636S.

Bradley EW, Ruan MM, Vrable A and Oursler MJ (2008) Pathway crosstalk between Ras/Raf and PI3K in promotion of M-CSF-induced MEK/ERK-mediated osteoclast survival. J Cell Biochem 104: 1439-1451.

Chen X, Sood S, Yang CS, Li N and Sun Z (2006) Five-lipoxygenase pathway of arachidonic acid metabolism in carcinogenesis and cancer chemoprevention. Current cancer drug targets 6: 613-622.

Colaprico A, Cava C, Bertoli G, Bontempi G and Castiglioni I (2015) Integrative Analysis with Monte Carlo Cross-Validation Reveals miRNAs Regulating Pathways Cross-Talk in Aggressive Breast Cancer. Biomed Res Int 2015: 831314.

Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, Caudy M, Garapati P, Gopinath G, Jassal B, et al. (2011) Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res 39: D691-697.

del Sol A, Balling R, Hood L and Galas D (2010) Diseases as network perturbations. Curr Opin Biotechnol 21: 566-571.

Huang H, Hara A, Homma T, Yonekawa Y and Ohgaki H (2005) Altered expression of immune defense genes in pilocytic astrocytomas. J Neuropathol Exp Neurol 64: 891-901.

Inoki K, Corradetti MN and Guan K-L (2005) Dysregulation of the TSC-mTOR pathway in human disease. Nature genetics 37: 19-24.

Ishii K, Zaitsu M, Yonemitsu N, Kan Y, Hamasaki Y and Matsuo M (2008) 5-lipoxygenase pathway promotes cell proliferation in human glioma cell lines. Clinical neuropathology 28: 445-452.

Jones DT, Gronych J, Lichter P, Witt O and Pfister SM (2012) MAPK pathway activation in pilocytic astrocytoma. Cell Mol Life Sci 69: 1799-1811.

Jones DT, Kocialkowski S, Liu L, Pearson DM, Bäcklund LM, Ichimura K and Collins VP (2008) Tandem duplication producing a novel oncogenic BRAF fusion gene defines the majority of pilocytic astrocytomas. Cancer research 68: 8673-8677.

Lambert SR, Witt H, Hovestadt V, Zucknick M, Kool M, Pearson DM, Korshunov A, Ryzhova M, Ichimura K, Jabado N, et al. (2013) Differential expression and methylation of brain developmental genes define location-specific subsets of pilocytic astrocytoma. Acta Neuropathol 126: 291-301.

Logan CY and Nusse R (2004) The Wnt signaling pathway in development and disease. Annu. Rev. Cell Dev. Biol. 20: 781-810.

Magoni D (2002) nem: A software for network topology analysis and modeling, Modeling, Analysis and Simulation of Computer and Telecommunications Systems, 2002. MASCOTS 2002. Proceedings. 10th IEEE International Symposium on: 364-371: IEEE.

Miller JA, Oldham MC and Geschwind DH (2008) A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. The Journal of neuroscience 28: 1410-1420.

Ono K and Han J (2000) The p38 signal transduction pathway activation and function. Cellular signalling 12: 1-13.

Scheithauer B, Hawkins C, Tihan T, VandenBerg S and Burger P (2007) Pilocytic astrocytoma. WHO classification of tumours of the central nervous system, 4th edn. IARC, Lyon14-21.

Tian M, Chen X, Xiong Q, Xiong J, Xiao C, Ge F, Yang F and Miao W (2014) Phosphoproteomic analysis of protein phosphorylation networks in Tetrahymena thermophila, a model single-celled organism. Molecular & Cellular Proteomics 13: 503-519.

von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA and Bork P (2005) STRING: known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic Acids Res 33: D433-437.

Ward E, DeSantis C, Robbins A, Kohler B and Jemal A (2014) Childhood and adolescent cancer statistics, 2014. CA Cancer J Clin 64: 83-103.

Wen PY and Kesari S (2008) Malignant gliomas in adults. N Engl J Med 359: 492-507.

Zhou X, Zheng R, Zhang H and He T (2015) Pathway crosstalk analysis of microarray gene expression profile in human hepatocellular carcinoma. Pathol Oncol Res 21: 563-569.

Zhuang DY, Jiang L, He QQ, Zhou P and Yue T (2015a) Identification of hub subnetwork based on topological features of genes in breast cancer. Int J Mol Med 35: 664-674.

Zhuang DY, Jiang L, He QQ, Zhou P and Yue T (2015b) Identification of hub subnetwork based on topological features of genes in breast cancer. Int J Mol Med 35: 664-674.