Network analysis reveals potential markers for pediatric adrenocortical carcinoma

Anurag Kulshrestha¹
Shikha Suman²
Rakesh Ranjan¹

¹Bioinformatics Division, National Bureau of Animal Genetic Resources, Karnal, ²Division of Applied Sciences, Indian Institute of Information Technology, Allahabad, India

Abstract: Pediatric adrenocortical carcinoma (ACC) is a rare malignancy with a poor outcome. Molecular mechanisms of pediatric ACC oncogenesis and advancement are not well understood. Accurate and timely diagnosis of the disease requires identification of new markers for pediatric ACC. Differentially expressed genes (DEGs) were identified from the gene expression profile of pediatric ACC and obtained from Gene Expression Omnibus. Gene Ontology functional and pathway enrichment analysis was implemented to recognize the functions of DEGs. A protein–protein interaction (PPI) and gene–gene functional interaction (GGI) network of DEGs was constructed. Hub gene detection and enrichment analysis of functional modules were performed. Furthermore, a gene regulatory network incorporating DEGs–microRNAs–transcription factors was constructed and analyzed. A total of 431 DEGs including 228 upregulated and 203 downregulated DEGs were screened. These genes were largely involved in cell cycle, steroid biosynthesis, and p53 signaling pathways. Upregulated genes, CDK1, CCNB1, CDC20, and BUB1B, were identified as the common hubs of PPI and GGI networks. All the four common hub genes were also part of modules of the PPI network. Moreover, all the four genes were also present in the largest module of GGI network. A gene regulatory network consisting of 82 microRNAs and 100 transcription factors was also constructed. CDK1, CCNB1, CDC20, and BUB1B may serve as potential biomarker of pediatric ACC and as potential targets for therapeutic approach, although experimental studies are required to authenticate our findings.

Keywords: gene expression profiling, protein–protein interaction network, network module, gene–gene functional interaction network

Introduction
Adrenocortical tumor (ACT) is an aberrant and highly aggressive malignancy originating from the adrenal cortex. It is accountable for ~0.2% of all childhood cancers. The majority of pediatric ACTs are functional as compared to adult ACTs, which are mostly nonfunctional.¹ Girls are more frequently affected with pediatric adrenocortical carcinoma (ACC) than boys.² Increased production of androgens, aldosterone, corticosteroids, and estrogen with ~80% showing virilization are the symptoms associated with pediatric ACC.³ Like most pediatric embryonal tumors, a good result demands early detection and complete surgical resection.⁴ Traditional chemotherapeutic agents have shown little value in treating children with ACC. Long-term problems associated with mitotane plus EDP (etoposide, doxorubicin, and cisplatin) treatment are a troublesome issue for children with ACC. Moreover, leukemogenesis may develop on treatment with topoisomerase inhibitors such as etoposide and doxorubicin.³ Surgery and exhaustive chemotherapy have shown poor outcomes in children with locally advanced or metastasis disease.⁴
Pediatric ACC is commonly reported in families with Li–Fraumeni syndrome, which are generally related with TP53 tumor-suppressor germ line mutations or inherent genetic and/or epigenetic changes affecting chromosome 11p15 (Beckwith–Wiedemann syndrome). Although the elements advancing to occasional pediatric ACCs are not established, their similarity to cases with an inherent susceptibility indicates a common method of tumorigenesis. Size of the tumor and verification of the tumor after surgery form the basis of staging of pediatric ACC. Early-stage tumor tissue can be accessed through the distinguishable clinical symptoms of ACC such as Cushing syndrome and virilization. Embryonal tumors share the epidemiological and molecular characteristics with ACC.

Molecular studies differentiating pediatric ACC from age-matched normal adrenals have been established from gene expression profiling. Rarity of this tumor has been a problem in identifying the potential markers. Increased levels of insulin-like growth factor 2 (IGF-2) are found in 90% of pediatric ACC due to genetic or epigenetic changes in chromosome 11p15. KCNQ1 (potassium channel, voltage-gated KQT-like subfamily Q, member 1) and CDKN1C (cyclin-dependent kinase inhibitor 1C) are among the most strongly downregulated genes in pediatric ACC. Genes associated with mitogen-activated kinase and growth factor receptor pathways are impaired in pediatric ACC, indicating their plausible utility as therapeutic targets. HSD3B2, a steroidogenic enzyme, and its transcriptional regulators NR4A1 (nuclear receptor subfamily 4, group A, member 1) and NR4A2 (nuclear receptor subfamily 4, group A, member 2) are downregulated in pediatric ACC. Another highly downregulated gene in pediatric ACC is NOV (nephroblastoma overexpressed), which encodes a multimodular protein that has proapoptotic function on adrenocortical cancer cells. MicroRNA (miRNA) expression profiling of pediatric ACC revealed downregulation of miR-99a and miR-100, which in turn downregulates the expression of insulin-like growth factor 1 receptor (IGF-1R), mechanistic target of rapamycin (mTOR), and regulatory associated protein of mTOR, complex 1 (RPTOR) in adrenocortical cell lines.

A listing of differentially expressed genes (DEGs) is provided through gene expression analysis. Protein–protein interactions (PPIs) utilize known relationships among the protein molecules. Analyzing the PPI network recounts the importance of these interactions in relation to signal transduction and biochemistry. In a particular biological context, all proteins interact with other proteins to perform particular functions. Substantial amount of attention has been given to coherent analysis of microarray gene expression data with PPI networks in recent years. Potential biomarker identification, elucidation of protein function and protein interaction, functional module identification, and drug target identification are some of the applications of analyzing PPI networks.

This study focuses on analyzing the gene expression profile of children with ACC, based on the understanding of interaction networks utilizing a system biology approach. To obtain more knowledge from gene expression profiles, analysis should transcend identification of the affected genes leading to the proteins underlying the inflated biological phenotypes.

Materials and methods
A bioinformatics approach with myriad of computational tools, software, and databases was utilized for shedding light on the underlying mechanisms of pediatric ACC. The entire workflow representing the procedure employed for the study is shown in Figure 1.

Raw biological data
The raw DNA microarray data were obtained from Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) for healthy children and children suffering from ACC. The chip dataset GSE75415 included samples from healthy children, children with adenomas, and children with ACC. Eighteen ACC and seven normal adrenal samples were extracted from the dataset. Studying the gene expression enables to identify potential biomarkers for early detection of ACC. Gene expression profiling was performed using Affymetrix Human Genome U133A chips. (Affymetrix, Inc., Santa Clara, CA, USA)

Data preprocessing and normalization
DNA microarray analysis begins with preprocessing and normalization of raw biological data. This process removes noise from the biological data and ensures its integrity. Background correction of probe data, normalization, and summarization were executed by robust multi-array average analysis algorithm in affy package of R. Raw intensity and normalized intensity box plot were also generated for the analysis.

Elucidation of DEGs
Normalization of data was followed by analysis of differential expression by Linear Models for MicroArray data package of R. Delineating parameters such as adjusted P-value, false discovery rate (FDR), and fold change were utilized for filtering of DEGs between healthy and diseased states. To reduce error from multiple hypothesis testing,
Benjamini–Hochberg method was employed to obtain adjusted P-values. Genes were further screened on account of adjusted P-value <0.05 and fold change >2.

**Functional enrichment of DEGs**

Discerning the implication of DEGs in ACC, biological attributes of DEGs such as biological processes, molecular functions, and cellular components were extracted from Gene Ontology (GO) enrichment analysis. Database for Annotation, Visualization, and Integrated Discovery, WEB-based Gene SeT AnaLysis Toolkit (WebGestalt), and Funrich tools were surveyed for GO and pathway enrichment of DEGs. Hypergeometric test was utilized to assess the functional enrichment against predefined functional categories through large genomic, proteomic, and genetic datasets with P-value <0.05 and gene count >2.

**PPI network construction**

The recent version of STRING v10 database was employed for construction of PPI network of proteins encoded by DEGs. STRING is a database of known and predicted PPIs. The interactions include direct (physical) and indirect (functional) associations. The interactions are procured from genomic context, high-throughput experiments, coexpression, and previous knowledge. Three PPI networks were constructed by mapping upregulated DEGs, downregulated DEGs, and total DEGs, respectively, to STRING database with confidence score >0.4. PPI networks were visualized and analyzed in Cytoscape software, which furnishes diverse plugins for multiple analyses. Cytoscape represents PPI networks as graphs with nodes illustrating proteins and edges depicting associated interactions. Hub protein nodes of the PPI network with connectivity degree >10 were identified.

**Network topology analysis**

Network Analyzer was employed to analyze the topological parameters of networks. Architecture of complex networks was examined with topological parameters such as clustering coefficient, centralization, density, and network diameter. Undirected edges were considered for all the networks. The number of directly connected neighbors of a node in a network was termed as degree of a node. Node degree distribution \( P(k) \) is termed as the number of nodes with a degree \( k \) for \( k=0, 1, 2, \ldots \). Power law of distribution of node degrees, one of the most crucial network topological characteristics, was analyzed. A line can be fitted on the node degree distribution data to visualize the pattern of their dependencies. Network Analyzer uses the least squares method and only the points with positive coordinate values are considered for fit. R-squared value \( (R^2) \), also known as the coefficient of determination, gives the proportion of variability in the dataset. When \( R^2 \) value is close to 1, the fit is considered to be good. Also, other network parameters were analyzed.
Module identification and enrichment analysis
Module identification is imperative as two interacting proteins have a higher probability of sharing the same function than two proteins not interacting. Molecular Complex Detection (MCODE)\(^29\) was used to find local dense subnetworks or modular clusters through vertex weighing by local neighborhood density and outward traversal from a locally dense protein node to isolate dense regions in PPI network. Module identification criteria included degree cutoff of 2, node score cutoff of 0.2, k-core of 2, and maximum depth of 100. Significant modules were identified with MCODE score \(4\) and nodes \(6\).

Biological significance of these predicted modules was inferred by BiNGO\(^30\) plugin of Cytoscape. GO enrichment was performed with \(P\)-value \(<0.05\) based upon hypergeometric test and corrected by Benjamini and Hochberg FDR. GO Slim is utilized by BiNGO for functional annotation.

Construction of gene–gene functional interaction network
Gene–gene interaction network incorporating up- and downregulated DEGs was constructed for identifying the functional interactions between DEGs. DEGs’ list was furnished to GeneMANIA\(^31\) that incorporates large functional association data such as coexpression data, colocalization data, physical interactions, shared protein domains, pathway, and genetic interactions. Twenty additional genes were added to the interaction network based on GO term (biological process)-based weighting and \textit{Homo sapiens} as the species to identify genes that may have been missed during the screening process. Hub genes were identified with connectivity degree \(>10\). MCODE was employed for identification of modules in the gene–gene functional interaction (GGI) network. BiNGO and Gene Set Enrichment Analysis\(^32\) were utilized to identify the GO terms and pathways associated with modules with FDR \(q\)-value below 0.05.

Construction of transcription factor–miRNA regulatory network
Genes must interact with and respond to an organism’s environment, as they independently cannot control the organism by themselves. Transcription factors (TFs) and miRNA regulate the gene expression at transcriptional and posttranscriptional levels. Information pertinent to TFs, miRNAs, and their respective target genes may help to better understand the intrinsic processes of pediatric ACC. Molecular Signatures Database v5.1 (MSigDB)\(^32\) was explored for the identification of TFs and miRNAs associated with DEGs with FDR \(q\)-value below 0.05. MSigDB incorporates annotated gene sets derived from a large variety of resources. A gene regulatory network incorporating DEGs, TFs, and miRNAs was constructed in Cytoscape.

Results
The comprehensive study focused on identifying and analyzing the genes, proteins, and probable patterns that are expressed in children with ACC, as compared to normal children. The ACC dataset was exposed to preprocessing and normalization in order to remove noise by robust multivariate analysis algorithm (Figure 2).

![Figure 2](https://www.dovepress.com/)

**Figure 2** Preprocessing and normalization of expression data.
**Notes:** (A) Raw intensity box plot. (B) Normalized intensity box plot. Bars represent interquartile range. Red in (A) represents dataset before normalization and blue (B) after normalization.
Markers of pediatric ACC

Identification of DEGs
The normalized data were subjected to analysis to reveal genes with altered expression profiles between healthy and diseased datasets. A total of 431 DEGs were obtained through a threshold of adjusted $P$-value <0.05 and a fold change $\geq$2. Among the total DEGs, 228 were upregulated and 203 were downregulated genes.

Functional enrichment of DEGs
Biological significance of DEGs was established by enriching the GO functions such as biological processes, cellular components, and molecular functions. Among the total DEGs, both upregulated and downregulated DEGs were largely involved in metabolic process and protein binding. Most of the upregulated DEGs were present on the nucleus, while the downregulated DEGs existed on the membrane (Figure 3).

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis revealed that the upregulated DEGs were mostly enriched in cell cycle, steroid biosynthesis, and p53 signaling pathway, while the downregulated DEGs were mainly involved in complement and coagulation cascades (Table 1). Funrich enrichment analysis of the biological pathways associated with DEGs is shown in Figure 4.

PPI network construction
STRING furnishes original reliable protein data for consequent analysis. Upregulated, downregulated, and total DEGs were mapped to generate three PPI networks. A PPI network was formed with upregulated DEGs containing 194 nodes and 1,122 edges, as these DEGs had literature related to interacting proteins. Moreover, a PPI network was formed with 130 nodes and 262 edges with the available literature. The total DEG PPI network was formed with 366 nodes and 1,858 edges (Figure 5A).

Hub genes of the total DEGs PPI network were identified as glyceraldehyde-3-phosphate dehydrogenase (GAPDH), cyclin-dependent kinase 1 (CDK1), topoisomerase (DNA) II alpha (TOP2A), cyclin B1 (CCNB1), and ATP citrate lyase (ACLY). Top 20 hub genes of the overall PPI network are shown in Table 2.

Network topological analysis
PPI networks or biological networks are notably different from random networks on the basis of differentiable topological characteristics. The most important indicator is the power law of node degree distribution. The power law of degree distribution was followed with an $R^2=0.749, 0.859,$
and 0.836 for upregulated, downregulated, and total DEGs, respectively. This implies that all the PPI subnetworks were scale-free, a major attribute of complex biological networks.\textsuperscript{33} Various parameters of the PPI networks such as clustering coefficient, network centralization, and network density are shown in Table 3.

### Module identification and enrichment analysis

The overall PPI network of the DEGs was surveyed for identification of functional modules in the network. Five modules were identified in the PPI network with MCODE score \(4\) and nodes \(6\): module P-A with MCODE score of \(8.625\) (nodes \(=17\)), module P-B with MCODE score of \(6.857\) (nodes \(=22\)), module P-C with MCODE score of \(5.727\) (nodes \(=23\)) (Figure 5B), module P-D with MCODE score of \(5.2\) (nodes \(=16\)) (Figure 5C), and module P-E with MCODE score of \(5\) (nodes \(=19\)). Hub genes, namely, CDK1, CCNB1, and cell division cycle 20 (CDC20), were present in module P-D, and BUB1 mitotic checkpoint serine/threonine kinase B (BUB1B) was present in module P-C. Modules P-C and P-D were scrutinized further for function and pathway enrichment analysis.

GO functional enrichment of the modules revealed that module P-C was enriched in sterol and steroid metabolic process and that module P-D was enriched in cell cycle process and cell cycle phase (Table 4). Also, KEGG pathway enrichment analysis revealed that genes in module P-C were enriched in cell cycle and p53 signaling pathway and that genes in module P-D were significantly enriched in aminoacyl-tRNA biosynthesis and cell cycle pathways (Table 5).

<p>| Table 1 KEGG pathway enrichment analysis of DEGs (top five in each) |</p>
<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Count</th>
<th>Adjusted (P)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upregulated</td>
<td>KEGG hsa04110 Cell cycle</td>
<td>20</td>
<td>2.70E-09</td>
</tr>
<tr>
<td></td>
<td>KEGG hsa00100 Steroid biosynthesis</td>
<td>6</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>KEGG hsa04110 p53 signaling pathway</td>
<td>9</td>
<td>0.0051</td>
</tr>
<tr>
<td></td>
<td>KEGG hsa04114 Oocyte meiosis</td>
<td>11</td>
<td>0.0049</td>
</tr>
<tr>
<td></td>
<td>KEGG hsa00900 Terpenoid backbone biosynthesis</td>
<td>4</td>
<td>0.0961</td>
</tr>
<tr>
<td>Downregulated</td>
<td>KEGG hsa04610 Complement and coagulation cascades</td>
<td>9</td>
<td>0.0012</td>
</tr>
<tr>
<td></td>
<td>KEGG hsa02010 ABC transporters</td>
<td>4</td>
<td>0.8510</td>
</tr>
</tbody>
</table>

**Abbreviations:** DEGs, differentially expressed genes; KEGG, Kyoto Encyclopedia of Genes and Genomes.

![Figure 4 Pathway enrichment analyses](https://www.dovepress.com/figure-4-pathway-enrichment-analyses-a-upregulated-and-b-downregulated-degs.png)

**Figure 4** Pathway enrichment analyses of (A) upregulated and (B) downregulated DEGs.

**Abbreviations:** DEGs, differentially expressed genes; IGF, insulin-like growth factor; IGFBPs, insulin-like growth factor binding proteins.
Figure 5 PPI network of pediatric ACC along with the modules.

Notes: (A) Overall PPI network of the DEGs. Upregulated genes are depicted in blue (circle) and downregulated genes in red (triangle). (B) Module with McODe score 5.2 and (C) module with McODE score 5.727.

Abbreviations: ACC, pediatric adrenocortical carcinoma; DEGs, differentially expressed genes; McODe, molecular complex detection; PPI, protein–protein interaction.
### Table 2: Top 20 hubs of overall PPi network

<table>
<thead>
<tr>
<th>ID</th>
<th>Degree</th>
<th>ID</th>
<th>Degree</th>
<th>ID</th>
<th>Degree</th>
<th>ID</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>88</td>
<td>CDC20</td>
<td>51</td>
<td>HSPAS5</td>
<td>38</td>
<td>RRM2</td>
<td>33</td>
</tr>
<tr>
<td>CDKI</td>
<td>76</td>
<td>AURKA</td>
<td>51</td>
<td>HDAC1</td>
<td>37</td>
<td>HSPA4</td>
<td>33</td>
</tr>
<tr>
<td>TOP2A</td>
<td>74</td>
<td>CCNB2</td>
<td>49</td>
<td>EGR1</td>
<td>36</td>
<td>SMARCA4</td>
<td>32</td>
</tr>
<tr>
<td>CCNB1</td>
<td>60</td>
<td>FOS</td>
<td>49</td>
<td>BUB1B</td>
<td>35</td>
<td>MMP2</td>
<td>32</td>
</tr>
<tr>
<td>ACY</td>
<td>52</td>
<td>CAT</td>
<td>40</td>
<td>MCM5</td>
<td>34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: DEGs, differentially expressed genes; PPi, protein–protein interaction.

### Table 3: Topological parameters of PPi networks

<table>
<thead>
<tr>
<th>PPI network</th>
<th>Number of nodes</th>
<th>Number of edges</th>
<th>$R^2$</th>
<th>Correlation</th>
<th>Clustering coefficient</th>
<th>Network centralization</th>
<th>Network density</th>
<th>Network diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall DEGs</td>
<td>366</td>
<td>1,858</td>
<td>0.836</td>
<td>0.880</td>
<td>0.304</td>
<td>0.214</td>
<td>0.028</td>
<td>9</td>
</tr>
<tr>
<td>Upregulated</td>
<td>194</td>
<td>1,122</td>
<td>0.749</td>
<td>0.849</td>
<td>0.404</td>
<td>0.280</td>
<td>0.060</td>
<td>8</td>
</tr>
<tr>
<td>Downregulated</td>
<td>130</td>
<td>262</td>
<td>0.859</td>
<td>0.996</td>
<td>0.213</td>
<td>0.204</td>
<td>0.031</td>
<td>9</td>
</tr>
</tbody>
</table>

Abbreviations: DEGs, differentially expressed genes; PPi, protein–protein interaction.

### Table 4: GO enrichment analysis of modules of PPI network (top five in each)

<table>
<thead>
<tr>
<th>Category</th>
<th>Class</th>
<th>Description</th>
<th>Gene count</th>
<th>Adjusted P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module P-C</td>
<td>BP</td>
<td>Sterol metabolic process</td>
<td>5</td>
<td>6.368E-05</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>Steroid metabolic process</td>
<td>6</td>
<td>6.368E-05</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>Amino acid activation</td>
<td>4</td>
<td>6.368E-05</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>tRNA aminoacylation</td>
<td>4</td>
<td>6.368E-05</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>tRNA aminoacylation for protein translation</td>
<td>4</td>
<td>6.368E-05</td>
</tr>
<tr>
<td>Module P-D</td>
<td>BP</td>
<td>Cell cycle process</td>
<td>9</td>
<td>1.6615E-06</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>Cell cycle phase</td>
<td>8</td>
<td>2.3258E-06</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>Cell cycle</td>
<td>9</td>
<td>8.3642E-06</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>Mitotic cell cycle</td>
<td>7</td>
<td>1.3412E-05</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>Organelle organization</td>
<td>10</td>
<td>3.8684E-05</td>
</tr>
</tbody>
</table>

Abbreviations: BP, biological process; GO, Gene Ontology; PPi, protein–protein interaction.

### Table 5: Pathway enrichment analysis of modules of PPI network (top five in each)

<table>
<thead>
<tr>
<th>Category</th>
<th>Class</th>
<th>Description</th>
<th>Gene count</th>
<th>Adjusted P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module P-C</td>
<td>KEGG</td>
<td>hsa04110 Cell cycle</td>
<td>6</td>
<td>6.13E-06</td>
</tr>
<tr>
<td></td>
<td>KEGG</td>
<td>hsa04115 p53 signaling pathway</td>
<td>4</td>
<td>8.55E-04</td>
</tr>
<tr>
<td></td>
<td>KEGG</td>
<td>hsa04114 Oocyte meiosis</td>
<td>4</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>KEGG</td>
<td>hsa04914 Progesterone-mediated oocyte maturation</td>
<td>3</td>
<td>0.0256</td>
</tr>
<tr>
<td>Module P-D</td>
<td>KEGG</td>
<td>hsa00970 Aminoacyl-tRNA biosynthesis</td>
<td>4</td>
<td>0.0030</td>
</tr>
<tr>
<td></td>
<td>KEGG</td>
<td>hsa04110 Cell cycle</td>
<td>3</td>
<td>0.3684</td>
</tr>
</tbody>
</table>

Abbreviations: PPI, protein–protein interaction; KEGG, Kyoto Encyclopedia of Genes and Genomes.
Markers of pediatric ACC

A GGI network of the overall DEGs was constructed to infer the biological meaning of the recognized DEGs at the gene level. The gene interaction network was composed of 449 nodes and 14,848 edges (Figure 6A). Approximately 64.66% genes show physical interactions, and 17.38% of genes show coexpression in the GGI network.

The hub genes of the GGI network were identified as ubiquitin C (UBC), interleukin enhancer binding factor 2 (ILF2), karyopherin subunit alpha 2, CCNB1, and CDC28 protein kinase regulatory subunit 2 (CKS2). The top 20 hubs of GGI network are shown in Table 6.

Ten modules were identified in the GGI network with MCODE score ≥4 and nodes ≥6. They were as follows: module G-A (MCODE score – 70.541) with 75 nodes (Figure 6B), module G-B (MCODE score – 18.698) with 64 nodes, module G-C (MCODE score – 12.963) with 28 nodes, module G-D (MCODE score – 7.333) with 28 nodes, module G-E (MCODE score – 5.647) with 35 nodes, module G-F (MCODE score – 5.556) with ten nodes, module G-G (MCODE score – 5.286) with 29 nodes, module G-H (MCODE score – 5.2) with eleven nodes, module G-I (MCODE score – 4.839) with 32 nodes, and module G-J (MCODE score – 4.833) with 13 nodes. Hub genes, namely, CDK1, CCNB1, CDC20, and BUB1B, were also present in the largest module G-A of the GGI network. Genes in module G-A were further scrutinized for enrichment analysis.

Genes in module G-A were found to be significantly enriched in cell cycle and cell cycle phase. Moreover, KEGG pathway enrichment revealed that genes in module G-A were enriched in cell cycle and oocyte meiosis pathways (Table 7).

Gene regulatory network

Identification of DEGs was preceded by recognizing TFs and miRNA associated with DEGs to better understand the process of gene regulation.

Eighty-two miRNAs such as TGGTGCT, MIR-29A, MIR-29B, MIR-29C, GTACTGT, MIR-101, GTGCCC, and MIR-506 were found to be associated with DEGs. One hundred TFs such as GGGCGGR_V$SP1_Q6 and TGGAAA_V$NFAT_Q4_01 were mapped to DEGs with FDR q-value below 0.05. Top ten TFs and miRNAs associated with DEGs are shown in Tables 8 and 9, respectively.

The DEGs–miRNAs–TFs regulatory network consisted of 520 nodes and 2,782 edges (Figure 7). Sp1 TF (SP1) was identified as the hub of the network with a node degree of 88.

Discussion

PPIs and protein expression are responsible for the pathological changes induced by the development of carcinoma. Multiple resources such as alterations in gene expression, PPI network, gene functional interaction network, hubs, and module identification were employed to identify potential diagnostic markers that can distinguish children with ACC.
A total of 431 DEGs including 228 upregulated and 203 downregulated DEGs were identified by microarray data analysis. Pathway enrichment analysis demonstrated that cell cycle, terpenoid backbone biosynthesis, and oocyte meiosis were overrepresented amid upregulated genes. IGF-2 was the most highly expressed gene in pediatric ACC, as compared to normal adrenal. Among the upregulated DEGs, CDK1, CCNB1, CDC20, and BUB1B were the common hubs among PPI and GGI networks. More centralized genes in the network are suggested to be key drivers of proper cellular function, in comparison to peripheral genes. Moreover, all these four genes were incorporated in the functional modules of the interaction networks.

CDK1, also known as CDC2, is a representative of serine/threonine protein kinase family. CDK1 is a catalytic subunit of M-phase promoting factor, a well-conserved protein kinase complex crucial for G1/S and G2/M phase transitions of cell cycle in eukaryotes. CDK1 has previously been reported to be upregulated in ACC samples by Glover et al. Moreover, in vivo inhibition of CDK1 by targeted miR-7 delivery has been proposed as a therapeutic approach for ACC. CDC2 was found to be dysregulated in cell cycle pathway through meta-analysis of gene expression and comprehensive genomic hybridization profiling data of ACC.

CCNB1 encodes for a regulatory protein involved in mitosis. Pinto et al reported that pediatric ACC based on TP53 and somatic ATRX mutations had shown high expression of genes associated with chromosome instability and deregulation of cell cycle control, such as CCNB1. Soon et al reported that CCNB1 expression was found to be appreciably higher in ACC as compared to adrenocortical adenoma. Also, the combination of IGF-2 and either MAD2L1, CCNB1, or Ki-67 is highly sensitive and specific for ACC.

CDC20 acts as regulatory protein involved in cell cycle progression, apoptosis, and ciliary disassembly. CDC20 has been proposed to exhibit oncogenic function, demonstrating its utility as a potential therapeutic target for combating human cancers. CDC20 has previously been proposed to be upregulated in transcriptome analysis of adrenocortical cancer.

BUB1B is an essential spindle assembly checkpoint protein that forms mitotic checkpoint complex, which on activation controls premature chromosome segregation. Mitotic inhibitor drugs such as taxanes, which disrupt the process of cell division, have proved to be potent anticancer drugs.

### Table 6 The top 20 hub genes of GGI network

<table>
<thead>
<tr>
<th>ID</th>
<th>Degree</th>
<th>ID</th>
<th>Degree</th>
<th>ID</th>
<th>Degree</th>
<th>ID</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBC</td>
<td>322</td>
<td>CDK2</td>
<td>135</td>
<td>H2AFZ</td>
<td>127</td>
<td>PLK1</td>
<td>123</td>
</tr>
<tr>
<td>IFL2</td>
<td>141</td>
<td>CKS1</td>
<td>133</td>
<td>CKS1B</td>
<td>127</td>
<td>PSMD14</td>
<td>122</td>
</tr>
<tr>
<td>KPNA2</td>
<td>138</td>
<td>PPC1G</td>
<td>131</td>
<td>UBE2S</td>
<td>125</td>
<td>DDX39A</td>
<td>122</td>
</tr>
<tr>
<td>CCNB1</td>
<td>137</td>
<td>BUB1B</td>
<td>130</td>
<td>MCM6</td>
<td>124</td>
<td>ZWNT4</td>
<td>121</td>
</tr>
<tr>
<td>CKS2</td>
<td>137</td>
<td>CDC20</td>
<td>127</td>
<td>DNM1T1</td>
<td>124</td>
<td>MCM3</td>
<td>121</td>
</tr>
</tbody>
</table>

**Abbreviation:** GGI, gene–gene functional interaction.
Combined expression of BUB1B–PINK1 has been proposed to be slightly related with disease-free survival in the pediatric group.41

A gene regulatory network incorporating DEGs–miRNAs–TFs was also constructed to better understand the process of gene regulation. Upon analysis, SP1 was found to be hub of the gene regulatory network. SP1 is a versatile sequence-specific DNA-binding protein involved in the expression of different genes.42 It is overexpressed in various human cancers and is involved in angiogenesis, cell growth, and resistance to apoptosis.43–45

The study has some limitations as the data utilized in the study consisted of 18 ACC and seven control samples, which

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Number of target genes</th>
<th>FDR q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGTTGCT, MIR-29A, MIR-29B, MIR-29C</td>
<td>20</td>
<td>1.72E-05</td>
</tr>
<tr>
<td>GACTGT, MIR-101</td>
<td>14</td>
<td>1.72E-05</td>
</tr>
<tr>
<td>GTGCTTT, MIR-506</td>
<td>21</td>
<td>3.41E-04</td>
</tr>
<tr>
<td>ATACCTC, MIR-202</td>
<td>10</td>
<td>3.41E-04</td>
</tr>
<tr>
<td>CAGTGGT, MIR-141, MIR-200A</td>
<td>13</td>
<td>3.41E-04</td>
</tr>
<tr>
<td>AATTGGA, MIR-23A, MIR-23B</td>
<td>15</td>
<td>3.86E-04</td>
</tr>
<tr>
<td>GTGACCT, MIR-224</td>
<td>9</td>
<td>4.89E-04</td>
</tr>
<tr>
<td>TGCCCTA, MIR-124A</td>
<td>17</td>
<td>4.89E-04</td>
</tr>
<tr>
<td>ACCAAAG, MIR-9</td>
<td>16</td>
<td>4.89E-04</td>
</tr>
<tr>
<td>ATGTGTC, MIR-494</td>
<td>9</td>
<td>4.89E-04</td>
</tr>
</tbody>
</table>

Abbreviations: DEGs, differentially expressed genes; FDR, false discovery rate; miRNA, microRNA.
were restricted in quantity and downloaded from the Gene Expression Omnibus database, not generated by us.

**Conclusion**

Four novel genes, CDK1, CCNB1, CDC20, and BUB1B, associated with pediatric ACC were identified by bioinformatics approaches. These DEGs were present in the hubs and modules of PPI and GGI networks, suggesting their potential utility as potential biomarker for pediatric ACC. These genes may also provide prospective targets for pediatric ACC therapy, although further experimental studies are essential to confirm the role of these genes and their potential to be developed as molecular targets for pediatric ACC.

**Acknowledgment**

Computing facilities at the National Bureau of Animal Genetic Resources, Karnal, India are gratefully acknowledged.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


