Co-expression network analysis of differentially expressed genes associated with metastasis in prolactin pituitary tumors

WEI ZHANG1, ZHENLE ZANG1, YECHUN SONG1, HUI YANG1 and QING YIN2

1Department of Neurosurgery, Xinqiao Hospital, Third Military Medical University, Chongqing 400037; 2Department of Rehabilitation and Physical Therapy, Southwest Hospital, Third Military Medical University, Chongqing 400038, P.R. China

Received August 23, 2013; Accepted March 5, 2014

DOI: 10.3892/mmr.2014.2152

Abstract. The aim of the present study was to construct a co-expression network of differently expressed genes (DEGs) in prolactin pituitary (PRL) tumor metastasis. The gene expression profile, GSE22812 was downloaded from the Gene Expression Omnibus database, and including five non-invasive, two invasive and six aggressive-invasive PRL tumor samples. Compared with non-invasive samples, DEGs were identified in invasive and aggressive-invasive samples using a limma package in R language. The expression values of DEGs were hierarchically clustered. Next, Gene Ontology (GO) function enrichment and Kyoto Encyclopedia of Genes and Genomes pathway analysis of DEGs were performed via The Database for Annotation, Visualization and Integrated Discovery. Finally, gene pairs of DEGs between non-invasive and aggressive-invasive samples were identified using the Spearman cor( ) function in R language. Compared with the non-invasive samples, 61 and 89 DEGs were obtained from invasive and aggressive-invasive samples, respectively. Cluster analysis showed that four genes were shared by the two samples, including upregulated solute carrier family 2, facilitated glucose transporter member 11 (SLC2A11) and teneurin transmembrane protein 1 (TENM1) and downregulated importin 7 (IP07) and chromogranin B (CHGB). In the invasive samples, the most significant GO terms responded to cyclic adenosine monophosphate and a glucocorticoid stimulus. However, this occurred in the cell cycle, and was in response to hormone stimulation in aggressive-invasive samples. The co-expression network of DEGs showed different gene pairs and modules, and SLC2A11 and CHGB occurred in two co-expression networks within different co-expressed pairs.

In the present study, the co-expression network was constructed using bioinformatics methods. SLC2A11, TENM1, IP07 and CHGB are hypothesized to be closely associated with metastasis of PRL. Furthermore, CHGB and SLC2A11 may be significant in PRL tumor progression and serve as molecular biomarkers for PRL tumors. However, further investigation is required to confirm the current results.

Introduction

Approximately 10% of all clinically overt intracranial neoplasms arise from the pituitary gland, of which prolactin-secreting adenomas, also termed prolactinomas, were the most common, accounting for 30-40% of pituitary tumors (1,2). Prevalence has classically been indicated to be 94 per 100,000 inhabitants (3). The early diagnosis of prolactin-secreting pituitary tumors is difficult due to the early signs or symptoms being associated with the overproduction of hormones, and hormones that are only combined with a single tumor are particularly rare (2,4). Hormones, such as estrogens, are apparently not the only cause and other possible hormones, including local growth factors in pituitary prolactinomas, are less well defined (5).

Prolactin pituitary (PRL) tumors are classified and managed according to size (6). For example, microprolactinomas (<1 cm) do not typically invade the parasellar region, while macroprolactinomas (>1 cm) are more likely to locally invade and compress surrounding structures. Previous studies have determined the molecular pathogenesis of pituitary tumor aggressive behavior and malignant transformation. A dopamine agonist was identified to control PRL via decreased D2 receptor availability, differential isoform expression and disrupted autocrine growth factor signaling (7,8). ErbB receptors, epidermal growth factor receptor (9), p185her2/neu (10), ErbB3 and ErbB4 (11), including the kinase-deficient ErbB3 dimerization with p185her2/neu may reflect tumor progression to an increasingly dedifferentiated state (12,13). Increased ErbB receptor expression has been verified in aggressive pituitary tumors and carcinomas. Furthermore, inhibition of the ErbB receptor may provide an alternative medical control of tumor growth and hormone secretion (14).

In the present study, the gene expression profile of PRL tumor samples, including non-invasive, invasive and aggressive-invasive samples, were investigated and the differentially
expressed genes (DEGs) associated with tumor metastasis were identified. The DEGs were subsequently clustered and The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to identify over-represented Gene Ontology (GO) categories in biological processes and significant pathways. Finally, gene pairs of DEGs between non-invasive and aggressive-invasive samples were identified and used to construct the co-expression network. The present study aimed to identify co-expressed gene pairs for PRL tumor metastasis via bioinformatics methods, with the purpose of investigating potential molecular biomarkers for use in clinical treatment.

Materials and methods

Microarray data and preprocessing. The gene expression profile GSE22812 (15) was downloaded from the Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/), and included 13 PRL tumor samples. These samples were classified as non-invasive (n=5), invasive (n=2) and aggressive-invasive (n=6). The annotation information of GeneChip was available based on the GPL2895 microarray platform (GE Healthcare/Amersham Biosciences CodeLink Human Whole Genome Bioarray).

The original probe-level data in the CEL files were converted into expression values with a log2 transformation. The probe signal was subsequently converted into the corresponding gene symbol using the microarray platform, GPL2895. For genes that corresponded with multiple probe sets, which exhibited a plurality of expression values, the expression values of those probe sets were averaged and the boxplot of the standardization expression value was mapped. Finally, these data were normalized using the between array normalization function in the limma package of R language (part of the Bioconductor project) (16,17).

DEGs analysis and screening. The limma package in R language was used to identify DEGs in the invasive and aggressive-invasive samples by comparison with the non-invasive samples. The DEGs with a fold change value (log2FC)>1 and P<0.05 were selected.

To avoid the difference of expression value in one group being greater than that between groups, the standard deviation (SD) of expression values of DEGs in the different groups was calculated. SD<0.15 served as the cut-off criterion for further DEG screening.

The expression values of DEGs were hierarchically clustered by Cluster software (http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm) (18) and the differences in gene expression between non-invasive, invasive, and aggressive-invasive samples were intuitively observed. The expression values of overlapping genes between invasive and aggressive-invasive samples were also hierarchically clustered.

Functional enrichment analysis of DEGs. The functional enrichment analysis for the screened DEGs from the invasive and aggressive-invasive samples was performed using DAVID (19) for Kyoto Encyclopedia of Genes and Genomes (KEGG) (20) pathway and GO (21) function analysis. P<0.01 served as the cut-off criterion. Comparison of the significant GO terms and KEGG pathways were used to identify the different biological functions between the invasive and aggressive-invasive samples.

Co-expression network analysis. The correlation coefficient (R value) of DEGs between non-invasive and aggressive-invasive samples was calculated via the Spearman method of cor( ) function in R language (22). The invasive samples were excluded due to the low number of samples. The linear association of the expression value between two genes was verified using the cor( ) test and only the gene pairs with |R|>0.95 and P<0.05 were selected to construct the interaction networks.

Results

Screening differentially expressed genes. The publicly available microarray dataset, GSE22812 was preprocessed and a total of 7,579 genes were obtained from 13 PRL tumor samples. The differences between samples were significantly reduced following normalization using the between array normalization function (Fig. 1). The median of the gene expression value was almost on a straight line, indicating a marked degree of standardization.

At log2FC>1, P<0.05 and SD<0.15, a total of 61 DEGs were obtained by comparing non-invasive and invasive samples, including 31 upregulated and 30 downregulated genes (Table I). A total of 89 DEGs were identified between non-invasive and aggressive-invasive samples, including 36 upregulated and 53 downregulated genes (Table II). The number of abnormally expressed genes from the aggressive-invasive samples was greater compared with the invasive samples.

Cluster analysis. Cluster analysis was used to determine the expression value of DEGs from the different samples. One aggressive-invasive sample (PRL_tumor_aggressive-invasive_6) showed a different expression value and was excluded. The other five aggressive-invasive, two invasive and five non-invasive samples were clustered and the color contrast is shown in Fig. 2. Briefly, a cluster analysis of five aggressive-invasive and five non-invasive samples showed two clusters; upregulated and downregulated genes in the aggressive-invasive samples (Fig. 2A). Similarly, cluster analysis of two invasive and five non-invasive samples showed two clusters; upregulated and downregulated genes in the invasive samples (Fig. 2B).

There were four genes that were shared by the invasive and aggressive-invasive samples, including solute carrier family 2, facilitated glucose transporter member 11 (SLC2A11), importin 7 (IPO7), chromogranin B (CHGB) and teneurin transmembrane protein 1 (TENM1). Cluster analysis showed upregulation of TENM1 and SLC2A11, and downregulation of IPO7 and CHGB in the aggressive-invasive samples (Fig. 2C).

Gene function annotation. To determine the function of DEGs in a PRL tumor, the DEGs were mapped into the GO database. Table III shows the significant GO terms in which the DEGs were primarily located. In the invasive samples, the majority of DEGs were enriched in response to cyclic adenosine monophosphate and a glucocorticoid stimulus. By contrast, in the aggressive-invasive samples, the cell cycle was the most significant GO term and the other genes partially correlated
Table I. Significant upregulated and downregulated differentially expressed genes (top 10 of each) from invasive samples.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Log_{2}FC</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDE4C</td>
<td>1.40</td>
<td>9.60E−04</td>
</tr>
<tr>
<td>PHF19</td>
<td>2.01</td>
<td>1.00E−03</td>
</tr>
<tr>
<td>SLC2A11</td>
<td>1.34</td>
<td>1.40E−03</td>
</tr>
<tr>
<td>KRTAP19-1</td>
<td>1.23</td>
<td>1.50E−03</td>
</tr>
<tr>
<td>ANKRD33B</td>
<td>2.09</td>
<td>1.60E−03</td>
</tr>
<tr>
<td>KLHL8</td>
<td>1.50</td>
<td>2.00E−03</td>
</tr>
<tr>
<td>GPR52</td>
<td>1.29</td>
<td>2.40E−03</td>
</tr>
<tr>
<td>EGFR-AS1</td>
<td>1.23</td>
<td>2.90E−03</td>
</tr>
<tr>
<td>PLCZ1</td>
<td>1.22</td>
<td>3.20E−03</td>
</tr>
<tr>
<td>ASGR1</td>
<td>1.06</td>
<td>4.40E−03</td>
</tr>
<tr>
<td>CDO1</td>
<td>-2.33</td>
<td>4.89E−05</td>
</tr>
<tr>
<td>TAGLN3</td>
<td>-1.46</td>
<td>4.65E−04</td>
</tr>
<tr>
<td>AKAP12</td>
<td>-2.33</td>
<td>1.30E−03</td>
</tr>
<tr>
<td>CORO1C</td>
<td>-1.45</td>
<td>2.30E−03</td>
</tr>
<tr>
<td>LAPT4M4B</td>
<td>-1.17</td>
<td>2.40E−03</td>
</tr>
<tr>
<td>JUNB</td>
<td>-1.46</td>
<td>3.30E−03</td>
</tr>
<tr>
<td>HPCAL1</td>
<td>-1.63</td>
<td>4.00E−03</td>
</tr>
<tr>
<td>ACTN1</td>
<td>-1.30</td>
<td>4.20E−03</td>
</tr>
<tr>
<td>HIST1H2AG</td>
<td>-1.45</td>
<td>5.70E−03</td>
</tr>
<tr>
<td>ATP1B3</td>
<td>-1.08</td>
<td>6.20E−03</td>
</tr>
</tbody>
</table>

Log_{2}FC, fold change value; PDE4C, phosphodiesterase 4C; cyclic adenosine monophosphate specific; PHF19, Plant Homeo Domain finger protein 19; SLC2A11, solute carrier family 2, facilitated glucose transporter member 11; KRTAP19-1, keratin associated protein 19-1; ANKRD33B, ankyrin repeat domain 33B; KLHL8, kelch-like family member 8; GPR52, G protein-coupled receptor 52; EGFR-AS1, epidermal growth factor receptor antisense RNA 1; PLCZ1, phospholipase C, ζ1; ASGR1, asialoglycoprotein receptor 1; CDO1, cysteine dioxygenase type 1; TAGLN3, transgelin 3; AKAP12, kinase (PRKA) anchor protein 12; CORO1C, coronin, actin binding protein, 1C; LAPT4M4B, lysosomal protein transmembrane 4B; JUNB, jun B proto-oncogene; HPCAL1, hipoocalcin-like 1; ACTN1, actinin, α 1; HIST1H2AG, histone cluster 1, H2ag; ATP1B3, adenosine triphosphatase, Na+/K+ transporting, β3 polypeptide.

with the response to hormone stimulation (Table III). DEGs from the aggressive-invasive samples were enriched into GO terms compared with invasive samples. However, DEGs were not significantly enriched into a specific pathway via the KEGG pathway analysis, which may be due to the limitation of the low number of DEGs that were analyzed.

Construction of a co-expression network. To construct the co-expression network of DEGs between non-invasive and aggressive-invasive samples, the correlation coefficient of DEGs was calculated and only gene pairs with |R|>0.95 and P<0.05 were selected. The co-expression network of DEGs from the invasive samples consisted of 42 nodes and 34 co-expressed pairs, presenting scattered gene pairs and only one module with more than five nodes (Fig. 3A). In the aggressive-invasive samples, the co-expression network of DEGs contained 68 nodes and 122 co-expressed pairs gathered together, forming positive correlation modules and mixed modules (Fig. 3B). There were five modules containing more than five nodes (Fig. 3B).

Briefly, SLC2A11 and CHGB were shared by two co-expression networks within the different co-expressed pairs. Furthermore, no co-expressed pair simultaneously occurred in two co-expression networks. A number of downregulated DEGs in the aggressive-invasive samples were gathered to form positive correlation modules (Fig. 3B).

Discussion

PRL-secreting adenomas were the most common (~40%) type of pituitary adenomas and an important cause of hypogonadism, infertility and osteoporosis, as well as central compressive effects (8,23,24). Medical therapy with dopamine agonists are highly effective in controlling the tumor mass in the majority
of cases, however, complicated situations, including dopamine agonist resistance, pregnancy and malignant prolactinomas, may require multi-therapies involving surgery, radiotherapy, or a combination of the two (25,26). Progress in elucidating the natural development of prolactinomas was beneficial for disease management (27). However, currently there is no method to accurately identify the invasive pituitary tumors that are most likely to metastasize, which would enable early treatment prior to progression to malignant prolactinomas.

In the current study, a total of 61 and 89 DEGs were identified by comparing non-invasive samples with invasive and aggressive-invasive samples, respectively. Cluster analysis showed that four overlapping genes, *SLC2A11*, *TENM1*, *IPO7* and *CHGB* were co-expressed, indicating the close association of these genes with tumor transformation. It has been reported that *CHGB* (*CgB*) may be considered as a universal granular marker for pituitary adenomas (28). In addition, the hypothalamic hormone, gonadotropin-releasing hormone and the CgA-derived peptide, pancreastatin, may regulate *CgB* mRNA in gonadotroph adenomas, indicating an autocrine effect of pancreastatin on pituitary tumor function (29). *CgB* is a precursor of two peptides, GAWK (*CHGB* 420-493) and BAM-1745 (1745-dalton pyroglutamyl) (30). Although the other three genes have not been verified to be associated with pituitary adenomas, previous studies have implicated these genes in other cancers. Notably, β-karyopherin genes (*IPO7*) were the importin-α/β complex that was frequently overexpressed in cancer and is
mediated by the nuclear importing of proteins with a classical nuclear localization signal (31,32). *SLC2A11*, the solute carrier 2A11 gene, encodes a novel sugar transporter, Human glucose transporter 11, and consists of 12 exons, located on chromosome 22q11.2. In human tissue, a 7.2-kb transcript of *SLC2A11* was detected exclusively in heart and skeletal muscle (33).
brates, there are four paragons, TEMPL 1-4, which are expressed prominently in the developing central nervous system. TEMPL 1 and -2 are expressed by distinct interconnected populations of neurons (34). Based on their distinct complementary expression, a possible function in the establishment of proper connectivity in the brain was hypothesized (35). Progress in elucidating the mechanisms underlying the pathogenesis of prolactinomas may enable future development of novel molecular therapies for treatment-resistant cases.

GO function enrichment analysis showed different GO terms of invasive and aggressive-invasive samples, indicating a different biological progression in the two stages. Hormone genes were abnormally expressed in the invasive samples and, in the aggressive-invasive samples, dysregulation of cell cycle progression was hypothesized as significant in malignant metastasis. In addition, the co-expression network of DEGs showed that SLC2A11 and CHGB occurred in two co-expression networks combined with different co-expressed pairs. These results indicated that CHGB and SLC2A11 may be significant in PRL tumor progression and may serve as molecular biomarkers for PRL tumors. However, samples were limited in the current study and co-expression of metastasis-associated genes was not comprehensively elucidated.

In conclusion, four genes that are relevant to tumor metastasis were identified from 13 PRL tumor samples using a bioinformatics method. Two gene pairs were shared by co-expression networks of DEGs from invasive and aggressive-invasive samples. These candidate factors correlated with the progression of prolactinoma and, thus, may provide a series of potential therapeutic targets for the treatment of pituitary tumors.

Acknowledgements

The present study was supported by a grant from the Natural Science Foundation Project of CQCSTC (grant no. 2010BB5179).

References