ORIGINAL ARTICLE

Heat Shock Protein 90 Alpha, Lass B Member 1: A Key Regulative Gene of Monitoring Duodenal Tumor for Obesity Induced by High-Fat Diet

Li Zheng 1, Jie Bai 2, Jin-Ling Zhang 1, Xia Meng 1, Yan-Na Hao 1

1 Department of Pharmacy, Xingtai Medical College, Xingtai, China
2 Department of Computer, Xingtai Medical College, Xingtai, China

SUMMARY

Background: Effective methods of preventing and treating duodenal carcinoma, especially cancer induced by obesity (resulting from a high-fat diet) remains a longstanding challenge in medicine.

Methods: With regard to the objective, key genes were explored in the evolutio nal process in a group of normal, obese (high-fat diet), and duodenal tumor mice. Here, 23 genes were selected by the bioinformatics method. In order to correct the result, verification experiments were performed twice through online analysis.

Results: Finally, heat shock protein 90 alpha class B member 1, enriched in inflammation, tumors and steroid hormones-related pathways, was the statistically different gene in the evolutional process.

Conclusions: This work provided a new perspective to understand the evolutional process in a group of normal, obese (high-fat diet), and duodenal tumor mice and a potential target gene for monitoring duodenal tumors for normal individuals especially for obesity induced by high-fat diet.


KEY WORDS

obesity, duodenal tumors, HSP90AB1, gene, biomarker

LIST OF ABBREVIATIONS

HSP90AB1 - heat shock protein 90 alpha, class B member 1
PDC - primary duodenal cancer
GC - gastrointestinal tumor
SI - small intestine
N-O-D - normal, obese (high-fat diet), and duodenal tumor mice
DEGs - differentially expressed genes
SAM - Significance Analysis of Microarrays
STC - series test of cluster
PPI - protein-protein interaction
GO - Gene Ontology
KEGG - Kyoto Encyclopaedia of Genes and Genomes
FDR - false discovery rate
NLR - NOD-like receptor
NUSE - Normalized Unscaled Standard Errors

Manuscript accepted April 3, 2017
RLE - Relative Log Expression

INTRODUCTION

Primary duodenal cancer (PDC), a rare gastrointestinal tumor (GC) [1], is a primary malignant tumor of the duodenum excluding the lower part of the common bile duct, ampulla of Vater, and pancreatic head tumor. Although the duodenum accounts for 1/10 of the total length of the small intestine (SI), PDC accounts for most of the SI malignant tumors. Therefore, more attention should be paid to PDC, an important kind of gastrointestinal malignancy.

Diagnosis of the PDC has remained elusive. Untypical clinical symptoms and low incidence add to the difficulty for diagnosis, which results in the high mortality of PDC. It is difficult to distinguish PDC from duodenal ulcer according to clinical signs and symptoms. Additionally, Buchbjerg T et al. report that the incidence of PDC is 5.4/1,000,000 in the southern part of Denmark [2]. In that area, 5-year survival (YR) of the PDC curative patients is 27% while that of PDC palliative patients is 0 [2]. In China, the median survival time of PDC curative patients is 49.0 months, while that of PDC palliative group is only 7.0 months [3]. Similar results show that though the incidence of PDC increases year by year, the overall incidence is still at a very low level (1997-2012), with 16% 5-YR [4]. All of those factors lead to poor prognosis.

Some people tried to explore risk factors of PDC to improve detection rate. The risk factors may be defined into two types: inflammation and non-inflammation. Risk factors associated with non-inflammation include stomach neoplasms [5] and colorectal neoplasms [5]. Risk factors related with inflammation are Crohn's disease and other inflammatory duodenal ulcers [6], upper gastrointestinal polyps [7], bowel disease caused by genetic abnormalities [8-9], and fat [10-18]. Obesity is the hot spot of the research due to its prevalence in the overall population.

Obesity plays an important role in PDC according to the Epidemiological Association of Duodenal Cancer. Though obesity is closely associated with GC in epidemiology, its mechanisms remain unclear to date [10]. The adipose tissue provides a more suitable microenvironment for tumor through some pro-inflammatory cytokines [11]. Since it stimulates the growth of stromal cell population, excessive fat deposition aggravates cancer [11]. Furthermore, the excessive fat aggravates inflammation in the gastrointestinal tract [10-12]. Finally, CYP1A-humanized mice (induced by high-fat diet) have small intestinal tumors after long-term feeding with 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine [13].

While the mechanism of the relationship of obesity and PDC remains unclear to date, Wang H et al. report that the abnormality of Apc occurs in obesity-induced small bowel cancer [13]. Stasi C et al. report that ghrelin, produced in the gastrointestinal tract, plays a contradictory role in obesity, and it can penetrate the blood-brain barrier [14]. On the one hand, ghrelin has a negative regulatory function on inflammatory cell factors, such as IL-6 and TNF-α, on the other hand, it has a positive regulatory function on appetite resulting from obesity and inflammation. Furthermore, it is reported that low visceral adiposity is related with poor prognosis among Japanese old people [15].

In addition, there are a few molecular mechanical studies between high-fat diet induced obesity and rare diseases, such as PDC. Some researchers have studied obesity induced by high-fat diet, since excessive intake of a high-fat diet is one of the main contributors to obesity. Schulz M D et al. believe the K-ras mutation (G12Dint) is the key gene in the microbial metabolism of small intestine tumor progress [16]. High-fat diet induced obesity leads to the regulatory disorder of glucose, lipid metabolism, and intestinal microbial metabolism [17]. To sum up, a hypothesis is proposed that key regulatory genes may exist in the evolutional process in a group of normal, obese (high-fat diet), and duodenal tumor mice (N-O-D). Efforts are focused on exploring the key genes in the process in a group of N-O-D. Moreover, the corresponding pathological changes are also examined, and the expressive data of GSE13298 and GSE42840 are downloaded. Subsequently, preliminary screening of different genes (DEGs) are performed with the method of biological information. Then, the DEGs are verified twice by the expressive data of GSE64640 and GSE69306. Lastly, HSP90AB1, the key gene in the process of N-O-D is described for the first time in our work. In conclusion, in this research, it is discovered that HSP90AB1 is the key regulatory gene in the pathological process in the group of N-O-D, which may be the potential target gene for the prevention, diagnosis, and treatment of PDC. The study provided new perspective on the pathological process in the group of N-O-D.

MATERIALS AND METHODS

Data source and controlled genes

Expressions data of GSE13298 (supported by National Institutes of Health Grant CA-084301 (to R.K.)) [18] and GSE42840 (supported by Deutsche Krebshilfe (107977)) [16] were downloaded from GEO based on the GPL 8321 platform ([Mouse430A_2] Affymetrix Mouse Genome 430A 2.0 Array). GSE42840 included two normal diet consumption and two high-fat diet consumption group samples, while GSE13298 includes three duodenal tumor samples. The normalized verification of the sample was carried out by Robust Multi-chip Averages as reported [19].

Analysis of different genes

Comparative analysis of significantly differentially expressed genes (DEGs) was performed with the method of Significance Analysis of Microarrays (SAM) [20].
order to control the false positive rate, q-value was calculated [21]. Subsequently, a volcano plot was constructed using the most significant 1500 DEGs [22].

**Series test of cluster**
The series test of cluster (STC) was performed to select DEGs at a logical sequence (normality-obesity-tumor). The transformed values (log₂ ratios of raw expression values) were calculated. The algorithm was based on the random variance model to determine the significant counts in the trend of the number of genes as reported. (p-value < 0.01 (corrected by FDR (false discovery rate))) [23]. The profiles were most typical and can represent the actual situation.

**Screening key genes**

**Hob genes in the protein-protein interaction network**
The protein-protein interaction (PPI) network of the DEGs of the selected profiles was constructed by STRING (http://www.string-db.org) [24]. The genes were selected if the combined_score > 0.4. Then the hob genes were picked out if the number of gene occurrences > 3.

**Genes enriched in gene ontology pathway**
The genes were screened by Gene Ontology (GO) pathway enrichment analysis (Biological function) by dint of DAVID (https://david.ncifcrf.gov) [25,26]. p-value < 0.05 (corrected by FDR) was set as the cutoff condition.

**Genes enriched in Kyoto Encyclopaedia of Genes and Genomes pathway**
The genes were selected by Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway enrichment analyses also using DAVID [25,26]. The p-value < 0.05 (corrected by FDR) was set as the cutoff condition.

**Overlapped genes**
The overlap genes among the PPI network, GO pathway enrichment and KEGG pathway enrichment were selected by BioVenn analyses online (http://www.cmbi.ru.nl/cdd/biovenn/index.php) [27].

**Verification**
The key genes were selected from the overlapped genes by verification twice. The first verification was the data in GSE64640 (wildtype control: GSM1576074, GSM1576086, GSM1576087; Lrig1null tumor: GSM1576085, GSM1576083, GSM1576081, GSM1576079, GSM1576077, GSM1576075), which was performed by GEO2R analysis of GEO online. The second data for verification by dint of similar method was GSE69306 (normal diet: GSM1697429, GSM1697434, GSM1697439, GSM1697445; high-fat diet: GSM1697451, GSM1697457, GSM1697463, GSM1697468, GSM1697474, GSM1697479). The measured site was duodenal in mice. The first comparison result was performed based on p-value < 0.05 (corrected by Benjamini and Hochberg). The second criterion was p-value < 0.05 without correction. Since the feeding time by high-fat diet was shorter in the second verification (GSE69306: 4 weeks) compared with the data source (GSE42840: 18 weeks), the standard was relaxed. The key genes were filtered through the two verifications by the common genes among the significant genes of the two verifications and the overlapped genes.

**RESULTS**

**Identification of DEGs**
Quality controls were shown in Figure 1. For dataset GSM1051229, GSM1051230, GSM1051225, GSM1051226, GSM335823, GSM335824, GSM335825 (in GSE13298 and GSE42840), a total of 1500 DEGs were identified among the two normal control, two high fat-diet induced, and three duodenal neoplasm group samples. The volcano plot constructed by the DEGs was shown in Figure 2.

**Series test of cluster**
Four significant profiles were identified among 16 gene expression profiles revealed by STC analysis of DEGs (Figure 3) (p-value < 0.01, corrected by FDR). A part of the model profiles was selected: profile 5 with 146 DEGs among the four significant expression profiles (Table 1).

**Screening key genes**

**Hob genes in the PPI network**
The PPI network including 146 DEGs in profile 5 was made up by STRING to explore the key genes. Finally, 46 hub genes (the above repeated 3 times) were drawn from the string_interactions table (shown in Table 1).

**Genes enriched in GO pathway**
Based on the p-value (Table 1), through GO analysis of the 146 DEGs in profile 5, it was discovered that 116 genes were enriched in 7 varying biological progresses.

**Genes enriched in KEGG pathway**
There were 27 genes enriched in 13 different biological progresses through KEGG analysis of the 146 DEGs in profile 5 (Table 1).

**Analysis of the overlapped genes**
Twenty-three significant common genes were included in the above analysis by Venn diagram (Table 1).

**Verification**
HSP90AB1, heat shock protein 90 alpha family class B member 1, was the significant key gene through verification (Table 2). The relative pathways were selected through the method of Genes enriched in Gene Ontology pathway and Genes enriched in Kyoto Encyclopaed-
Table 1. The analysis of 146 significant genes in profile 5.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Number of alternative genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hob genes in PPI network by string</td>
<td>46</td>
</tr>
<tr>
<td>Enriched genes by GO pathway analyses</td>
<td>116</td>
</tr>
<tr>
<td>Enriched genes by KEGG pathway analyses</td>
<td>27</td>
</tr>
<tr>
<td>Overlap genes by Venn</td>
<td>23</td>
</tr>
</tbody>
</table>

PPI - protein-protein interaction, GO - Gene Ontology, KEGG - Kyoto Encyclopaedia of Genes and Genomes.

Table 2. Two verifications of the overlapped genes.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Number of alternative genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlapped genes by Venn</td>
<td>23</td>
</tr>
<tr>
<td>Genes filtered by first verification</td>
<td>8</td>
</tr>
<tr>
<td>Genes filtered by second verification</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3. The pathways HSP90AB1 enriched in by GO and KEGG pathway analyses.

<table>
<thead>
<tr>
<th>No.</th>
<th>GO</th>
<th>KEGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0070062:extracellular exosome</td>
<td>5200:paths in cancer</td>
</tr>
<tr>
<td>2</td>
<td>0005515:protein binding</td>
<td>4151:PI3K-Akt signaling pathway</td>
</tr>
<tr>
<td>3</td>
<td>0016020:membrane</td>
<td>4915:estrogen signaling pathway</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>4621:NOD-like receptor signaling pathway</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>4914:progesterone-mediated oocyte maturation</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>5215:prostate cancer</td>
</tr>
</tbody>
</table>

GO - Gene Ontology, KEGG - Kyoto Encyclopaedia of Genes and Genomes.

dia of Genes and Genomes pathway. The results were shown in Table 3.

DISCUSSION

In this work, key genes were explored in the assumed process among the evolutional process in a group of normal, obese (high-fat diet), and duodenal tumor mice by PPI, GO enriched, and KEGG pathways in the duodenum site of mice. The result indicated that HSP90AB1 was the key gene for preventing or diagnosing duodenal cancer among obese or normal individuals. In this study, it is found that the expression of HSP90AB1 increased in obesity and duodenal cancer compared with normal mice. The pathways and the enrichment channel of HSP90AB1 can be divided into several parts. Firstly, the positioning and role of HSP90AB1. The pathway 0070062 (extracellular exosome) indicated that HSP90AB1 might be an extracellular protein, while the pathway 0005515 (protein binding) and 0016020 (membrane) indicated that HSP90AB1 might be a cytosolic protein which was dependent of interaction with cell membranes. Secondly, the inflammation related role of HSP90AB1 indicated by the pathway 4621 (NOD-like receptor (NLR) signaling pathway), for the maturation and pyrophosphorylation of pro-inflammatory cytokines were regulated in the pathway 4621 by activating of caspase-1 induced by indicated NLR [28]. Thirdly, the tumor related roles of HSP90AB1. The pathway 4621 (NLR signaling pathway) [28] and 4151 (PI3K-Akt signaling pathway) were related to tumor apoptosis. Similarly, the pathway 5200
Figure 1. The median of the Normalized Unscaled Standard Errors (NUSE) value per chip and the median of the Relative Log Expression (RLE) values were criteria for the feasibility and reliability of the experimental design.

A: NUSE. The abscissa represents gene expressions in samples and the ordinate represents the NUSE values. Eligibility criteria: (1 - 0.2) < the median of NUSE values < (1 + 0.2). The higher the NUSE is consistent, the less the analysis is affected by the chip technology factor or the sample biological factors is. B: RLE. The abscissa was similar to the abscissa in A, while the ordinate represents relative RLE values. Eligibility criteria: -0.25 < the median of RLE value < 0.25. Similar to the NUSE, the higher the RLE is consistent, the more reliable the follow-up analysis results is.

Figure 2. The volcano plot is constructed by 1500 DEGs.

The abscissa represents gene expressions in samples and the ordinate represents relative fold changes. The genes are selected by the criterion of fold changes (-1.8 to +1.8). The closer it is to red, the higher the gene expression after normalization correction is. The closer it is to green, the lower the gene expression after normalization correction is.
Figure 3. One significant model profile by analysis of the gene expression patterns.

A. There are 16 boxes representing 16 model expression profiles. Each p-value summarizes the different genes in a model profile. In brief, four significantly different model profiles are labeled with asterisk (p-value < 0.01). Profile 5 belongs to those kinds of genes that upregulate first and then downregulate. B. Expression profile 5 increases first and then decreases during the development of disease. The horizontal axis represents normality (1), obesity (2) and tumor (3) phases while the vertical axis represents the expressions of normalized different genes. Each line represents a gene.

(pathways in cancer) and 5215 (prostate cancer) were transferred to cancer. Fourthly, steroid hormones were related with HSP90AB1 indicated by 4915 (estrogen signaling pathway) and 4914 (progesterone-mediated oocyte maturation).

The results were consistent with previous studies. HSP90AB1 (90kDa), a molecular chaperone mainly located in cytoplasm, is one of the family of HSP90 proteins with the full name of heat shock protein 90 alpha class B member 1 [29]. A large number of oncoproteins are regulated by the protein, such as CDK [29] and others. Cdk14 was the significant gene in the verification of normal and cancer in our study. Xu C et al. [30] reported that increased HSP90AB1 was closely related with the C allele of rs2282151 in Chinese colorectal cancers compared with normal tissues. Furthermore, since the protein is always in extracellular cytoplasm [29], it is an ideal biomarker. HSP90AB1 plays an important role in cancer treatment as a drug for targeted combination therapy because it is one of the key regulatory genes in multiple pathways related to cancer [29]. Additionally, HSP90AB1 may play an important role in obesity through BMAL1. On the one hand, the expression of BMAL1 is in positive correlation with HSP90 in the mammalian core clock [31], and on the other hand, there is a positive correlation between the methylation in some regions of BMAL1 and obesity [32]. Though the results of this study reveal a close relationship between obesity and duodenal cancer in mice, it is not as tight as expected. Only one gene was explored without duodenum-specific gene. Furthermore, some significant regulatory genes have low expression in normal, high expression level in obesity induced by high-fat diet, and low expression level in duodenal cancer. The location of the tumor may be the explanation. The duodenum is the site where pancreas and gallbladder secrete fat-related digestive juice is located. So, duodenal cancer may lead to lipid-metabolism disorder.

Interestingly, there were differences between the trend of changes in the screening and the verification. The former increases and then decreases, while the latter increases all the time. Different statistical methods may be the reason, for there were three samples in the first statistics, while two samples in the second statistics were counted twice.

The results suggested that it was possible to detect the expression level of HSPAB1 to monitor PDC among obese individuals who are at high risk of suffering from
tumor. Furthermore, HSPAB1 may be the potential gene for treatment of PDC. Additionally, the gene may be a screening criteria for dietary supplements in obesity. It should be noted that this study has only examined significant DEGs in the process in the group of N-O-D. More experiments should be conducted in cells and animals to verify the result that the duodenum lacks a specific gene.

CONCLUSION

To sum up, this work suggests that HSPAB1 is the key gene in the evolutionary process in the group of N-O-D. HSPAB1, as a molecular chaperone, plays an important role in the corresponding pathology with the changes in inflammation, tumors, and steroid hormones. Additionally, HSPAB1 may be a potential biomarker, target gene, and a criteria gene for screening dietary supplements for PDC, especially for obesity. However, further research is required since current samples are far from being sufficient.

Acknowledgement:
The authors would like to acknowledge the technical assistance of the online analysis platform GCBI website.

Ethical Approval:
This article does not contain any studies with human participants or animals performed by any of the authors.

Declaration of Interest:
The author declares that there is no potential competitive conflict of interest.

References:


