Abstract. – OBJECTIVE: The mechanism of gastroesophageal reflux disease (GERD) has gradually been understood at the molecular biological level, and acid is considered as the major cause of GERD. The aim of this study was to explore the molecular mechanism of GERD related with low pH by investigating the differential gene expression of oesophageal cells stimulated under a low pH environment for different time.

MATERIALS AND METHODS: Comparisons were made between the control normal samples (stimulated for 0 min) and low pH treat samples for various time points, and differentially expressed genes (DEGs) were identified, further bioinformatics analysis were carried out, including functional annotation and PPI network construction.

RESULTS: The result indicated that the number of DEGs was increased along with the time of acid exposure, and the EGR1, JUN and FOS were found in all enriched Gene Ontology terms with association scores between them was high.

CONCLUSIONS: All results suggested that EGR1, JUN, FOS may play important roles in the development of GERD. In a word, our results may reveal the contribution of gastric acid to GERD and the pathogenesis of GERD.

Key Words: Gastroesophageal reflux disease, Acid exposure stimulation; Differentially expressed gene (DEGs), Interaction network, Functional enrichment analysis.

Introduction

Gastroesophageal reflux disease (GERD) is a chronic relapsing acid-peptic disorder, results from reflux of the stomach and duodenum contents into esophagus\(^1,2\). GERD affects 20-30\% of the population in Western countries and is one of the most common clinical problems in daily practice\(^3\). The quality of life in patients with GERD was significantly affected because of the complications, including esophageal stricture, Barrett esophagus, and esophageal adenocarcinoma\(^4\). Acid is considered as the major cause of GERD\(^5\), and according to the pH of the refluxed material, GERD can be divided into acid (pH < 4), weekly acid (pH 4-7) and weakly alkaline (pH > 7)\(^6\). What more, it is reported that weakly acid reflux episodes more than acid reflux episodes causes symptoms in patients\(^7\). Therefore, it is of great importance to study the response of oesophageal cells under low pH (weakly acid) environments.

The mechanism of GERD has gradually been understood at the molecular biological level. Pro-inflammatory factors, such as inflammatory cytokines (interleukin-6 and -8), leukocyte infiltration and oxidative stress, have been demonstrated to be involved in the development of GERD\(^3,8-10\). Low pH environment has been suggested to induce the DNA fragmentation and apoptosis in tumor cells\(^11,12\). Duggan et al\(^13\) had utilized a transcriptomic and bioinformatics approach to assess regulation of gene expressions in response to low pH. In this study, for the elucidate the pathogenesis mechanism of weakly acid reflux, the gene expression profile of GSE2144 of oesophageal cell line SKGT4 samples exposed to low pH (pH 6.5) for different time course was further used. Beside the differentially expressed genes (DEGs) screening, the significant changed functions of DEGs at different time points were enriched, and the protein-protein interaction (PPI) network of the common DEGs of different times was constructed, followed by the functional annotations of genes in the network. Our study may also aid in understanding the impact of gastric acid on the progression of GERD.
**Materials and Methods**

**Affymetrix Microarray Data**

All persons have given their informed consent prior to their inclusion in the study, and all human studies have been approved by China Ethics Committee and performed in accordance with the ethical standards. The gene expression profile of GSE214413 was downloaded from GEO (http://www.ncbi.nlm.nih.gov/geo/) database and a total of ten samples were obtained. In more detail, there were two control samples of normal human tissues without acid stimulation (0 min), and eight samples of oesophageal cell lines SKGT4 exposed to low pH 6.5 environments for 30 min, 120 min, 180 min and 240 min, with two samples at each time course. The annotation information of the dataset was also downloaded based on the GPL96 platform (Affymetrix Human Genome U133A Array).

**Data Preprocessing and DEGs Selection**

Firstly, the raw data downloaded were converted into identifiable expression form and the missing data were supplemented14. Secondly, the complemented data were performed normalization15. Next, comparisons were made between control samples and low pH stimulated samples using the LIMMA (Linear Models for Microarray Data) package in R language, to identify the DEGs16. Genes only with p-value < 0.05 and llog fold change (FC) > 1 were selected as DEGs.

**Hierarchical Clustering Analysis**

Hierarchical clustering is a method of cluster analysis which seeks to build a hierarchy of clusters17. Based on that gene expression is time specific, hierarchical cluster analysis17 was used to observe the dynamic changes of gene expression at each time point.

**Functional Enrichment Analysis of DEGs at Different Time Points**

In order to find out the changes of biological functions of SKGT4 cells under the low pH exposure, functional enrichment analysis for the selected DEGs were conducted. DAVID (Database for Annotation, Visualization, and Integrated Discovery) bioinformatics resources consist of an integrated biological knowledge base and analytic tools aimed at systematically extracting biological functions from large gene or protein lists18. Over-presented functions of DEGs with the false discovery rate (FDR) < 0.05 were screened.

**PPI Network Construction**

The STRING (Search Tool for the Retrieval of Interacting Genes) database provides both exper-

<table>
<thead>
<tr>
<th>GO-ID</th>
<th>FDR</th>
<th>Description</th>
<th>Genes</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>48522</td>
<td>5.1945E-08</td>
<td>Positive regulation of cellular process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>48518</td>
<td>1.1575E-07</td>
<td>Positive regulation of biological process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>42221</td>
<td>0.000016473</td>
<td>Response to chemical stimulus</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>32501</td>
<td>0.010568</td>
<td>Multicellular organismal process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>50896</td>
<td>0.00014875</td>
<td>Response to stimulus</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>10556</td>
<td>0.0001055</td>
<td>Regulation of macromolecule biosynthetic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>9889</td>
<td>0.00013928</td>
<td>Regulation of biosynthetic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>60255</td>
<td>0.00031548</td>
<td>Regulation of macromolecule metabolic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>80090</td>
<td>0.00013819</td>
<td>Regulation of primary metabolic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>31323</td>
<td>0.00018956</td>
<td>Regulation of cellular metabolic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>19222</td>
<td>0.00029354</td>
<td>Regulation of metabolic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>19219</td>
<td>0.00047586</td>
<td>Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>31326</td>
<td>0.00047915</td>
<td>Regulation of cellular biosynthetic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>51171</td>
<td>0.00047985</td>
<td>Regulation of nitrogen compound metabolic process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>50794</td>
<td>0.00038703</td>
<td>Regulation of cellular process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>50789</td>
<td>0.00078033</td>
<td>Regulation of biological process</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>65007</td>
<td>0.0014324</td>
<td>Biological regulation</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
<tr>
<td>10468</td>
<td>0.00037962</td>
<td>Regulation of gene expression</td>
<td>EGR1</td>
<td>PIM1</td>
</tr>
</tbody>
</table>

Table 1. The results of functional enrichment of DEGs in PPI network.
Hierarchical Clustering Analysis

Dynamic expression changes of all the samples were analyzed using Cluster software. As shown in Figure 1B, compared with expression values of genes in the control samples, the gene expressions in the samples that under low pH exposure stimulation for 30 min, 120 min, 180 min, and 240 min were significant changed, especially for the samples treated for 240 min (Figure 1B).

Functional Enrichment Analysis of DEGs at Different Time Points

Respectively, 12, 12, 14, and 10 significant GO (Gene Ontology) terms with FDR < 0.05 under low pH exposure stimulation for 30 min, 120 min, 180 min and 240 min were enriched (Figure 2). Interestingly, these screened DEGs from each time point were all closely related with functions related with cell regulation.

PPI Network Construction

From the screened DEGs at different time points, 28 common were identified. The interactions of the common DEGs were searched using String (Figure 3A). PPI network was constructed using 73 interaction pairs with the confidence score larger than 0.4 (Figure 3B). In additions, interactions of the DEGs with score larger than 0.9 are shown in Table I, such as EGR1, JUN and FOS.
Figure 2. The significant functions of DEG from 4 time points: A, 30 min; B, 120 min; C, 180 min; and D, 240 min.
A total of 18 significant GO terms with FDR < 0.05 were enriched (Table II) of the genes in the PPI network. The most significant GO term was positive regulation of cellular process (FDR = 5.1945E-08). Furthermore, EGR1, JUN, FOS were found in every GO term.

**Discussion**

In order to better understand the molecular mechanism of GERD under low pH, the publicly available microarray dataset GSE2144 was obtained for bioinformatics analysis. In this study, a total of 285, 365, 277 and 421 DEGs were identified from the cells under low pH stimulation for 30 min, 120 min, 180 min and 240 min, respectively. Furthermore, combined with the result of cluster analysis, we inferred that the number of DEGs may increase with the extension of stimulation time. The result was in accordance with previous study that the extension of acid exposure time and the increasing of exposure extent were the main mechanisms of GERD disease. In addition, we constructed PPI network which contained 73 pairs of interaction among 20 common DEGs. Furthermore, EGR1, JUN, FOS were found in all enriched GO terms (Table II) of genes in the PPI network, meanwhile, with high interactional scores.

EGR1 (Early Growth Response protein 1) protein also known as Zif268 (zinc finger protein 225) or NGFI-A (nerve growth factor-induced protein A), belongs to the product of EGR gene family, which is located in the nucleus of the zinc finger

<table>
<thead>
<tr>
<th>Node1</th>
<th>Node2</th>
<th>Association score</th>
</tr>
</thead>
<tbody>
<tr>
<td>JUN</td>
<td>ATF3</td>
<td>0.999</td>
</tr>
<tr>
<td>JUN</td>
<td>FOS</td>
<td>0.999</td>
</tr>
<tr>
<td>LIFR</td>
<td>LIF</td>
<td>0.999</td>
</tr>
<tr>
<td>JUN</td>
<td>FOFL1</td>
<td>0.998</td>
</tr>
<tr>
<td>FOS</td>
<td>EGR1</td>
<td>0.997</td>
</tr>
<tr>
<td>EDN1</td>
<td>FOS</td>
<td>0.995</td>
</tr>
<tr>
<td>EDN1</td>
<td>JUN</td>
<td>0.994</td>
</tr>
<tr>
<td>JUN</td>
<td>FOSB</td>
<td>0.993</td>
</tr>
<tr>
<td>JUN</td>
<td>DUSP1</td>
<td>0.983</td>
</tr>
<tr>
<td>FOSB</td>
<td>EGR1</td>
<td>0.976</td>
</tr>
<tr>
<td>NR4A1</td>
<td>EGR1</td>
<td>0.976</td>
</tr>
<tr>
<td>NR4A2</td>
<td>NR4A1</td>
<td>0.972</td>
</tr>
<tr>
<td>FOSB</td>
<td>NR4A1</td>
<td>0.97</td>
</tr>
<tr>
<td>FOS</td>
<td>DUSP1</td>
<td>0.969</td>
</tr>
<tr>
<td>ATF3</td>
<td>EGR1</td>
<td>0.966</td>
</tr>
<tr>
<td>EGR1</td>
<td>DUSP1</td>
<td>0.965</td>
</tr>
<tr>
<td>JUN</td>
<td>NR4A1</td>
<td>0.958</td>
</tr>
<tr>
<td>JUN</td>
<td>EGR1</td>
<td>0.955</td>
</tr>
<tr>
<td>EDN1</td>
<td>CTGF</td>
<td>0.949</td>
</tr>
<tr>
<td>NR4A2</td>
<td>FOSB</td>
<td>0.944</td>
</tr>
<tr>
<td>FOSB</td>
<td>DUSP1</td>
<td>0.908</td>
</tr>
<tr>
<td>MAFF</td>
<td>ATF3</td>
<td>0.9</td>
</tr>
</tbody>
</table>

PPI: protein-protein interaction network.

*Functions Enrichment Analysis of DEGs in PPI Network*

A total of 18 significant GO terms with FDR < 0.05 were enriched (Table II) of the genes in the PPI network. The most significant GO term was
proteins and as a transcription factor for activates target genes related cell division and mitosis. Recent studies have shown that the expression level of EGR1 is changed in lesions and cancer tissues, suggesting that the EGR1 is involved in the development of cancers such as esophageal cancer, colon cancer and breast cancer. In addition, the EGR1 protein plays a pivotal role in the regulation of cell growth, differentiation and apoptosis. Lekkovitz et al. have reported that the EGR1 suppresses cerebellar granule cell apoptosis by blocking c-Jun activation. Meanwhile, Chen et al. identify EGR1 as a novel target for JUN-induced apoptosis in multiple myeloma. However, Hoffmann et al. consider JUN as an essential effector of EGR1 transcriptional regulation in inflammatory processes. Interestingly, the interaction between EGR1 and JUN was also observed in this study, thus, we inferred that the EGR/JUN complex may participate in the development of GERD or cancers via the interaction between EGR1 and JUN.

The FOS gene exhibits both oncogenic and tumor-suppressive functions, depending on the cellular context. FOS over-expression enhances the motility and invasion of breast and colorectal cancer cells, but inhibits the tumourigenicity of cervical carcinoma cell lines. Members of the FOS family (c-Fos, FosB and its smaller splice variants, Fra-1 and Fra-2) dimerise with JUN proteins to form the AP-1 transcription factor complex. All AP-1 complexes are characterized by a basic leucine-zipper region for dimerisation and DNA-binding, which have been implicated in transformation and progression of cancer soon after discovery. Interestingly, in this study, the interaction between JUN and FOS was found in PPI network, which indicated that our study was consist with the previous reports that some interaction was found between JUN and FOS.

In our work, FOS was interacted with JUN and EGR1. However, few researches have indicated the interaction between FOS and EGR1. Except that McMahon et al. suggested that EGR1 and FOS are co-regulated in some tissues. Therefore, our studies may provide some novel information for the interaction between FOS and EGR1, but more experiments are needed to support it.

Conclusions

In summary, based on the gene expression profile of GSE2144, the DEGs of oesophageal cells at low pH stimulation at various times were identified. Meanwhile, function enrichment analysis and PPI network were carried out, suggesting that interactions among EGR1, JUN and FOS play important roles in the progression of GERD, mainly by involving the positive regulation of cellular process and positive regulation of biological process. In a word, our studies may provide new ideas in investigate the molecular mechanism of GERD related with low PH. However, further experimental investigation are needed because our work is based on gene chips from a small sample size.

Conflict of Interest

The Authors declare that they have no conflict of interests.

References


13. Duggan SP, Gallagher WM, Fox EJ, Abdel-Latif MM, Reynolds JV, Kellerer D. Low pH induces co-ordin-


15. Fujita A, Sato J, Rodrigues L, Ferreira C, Sogayar M. Evaluating different methods of microarray da-


17. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide ex-

18. Da Wei Huang BTS, Lempicki RA. Systematic and integrative analysis of large gene lists using 

19. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Mingeux P, Doerks T, Stark T, Muller J, 


24. Patton WF. Fluorescence detection of glycopro-

25. Wu MY, Zhuiang CX, Yang HK, Liang YR. Expression of Egr-1, c-fos and cyclin D1 in esophageal 


28. Levkovitz Y, Baraban JM. A dominant negative inhibitor of the Egr family of transcription regulatory 


30. Hoffmann E, Ashour I, Wolter S, Doerrre A, Dif-


34. Tkach V, Tuchinsky E, Lukandin E, Vinson C, Bock E, Berezin V. Role of the Fos family members, c-


36. McMahon AP, Champion J, McMahon J, Sukhatme V. Developmental expression of the putative trans-

936