

# Analyzed gene expression of oesophageal cells stimulated in a low pH environment by DNA microarray

H.-X. YAO, W.-X. CHEN

Department of Gastroenterology, Shanghai JiaoTong University Affiliated Sixth People's Hospital, Shanghai, China

**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<sup>1-2</sup>. GERD affects 20-30% of the population in Western countries and is one of the most common clinical problems in

daily practice<sup>3</sup>. The quality of life in patients with GERD was significantly affected because of the complications, including esophageal stricture, Barrett esophagus, and esophageal adenocarcinoma<sup>4</sup>. Acid is considered as the major cause of GERD<sup>5</sup>, and according to the pH of the refluxed material, GERD can be divided into acid (pH < 4), weakly acid (pH 4-7) and weakly alkaline (pH > 7)<sup>6</sup>. What's more, it is reported that weakly acid reflux episodes more than acid reflux episodes causes symptoms in patients<sup>7</sup>. 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<sup>3,8-10</sup>. Low pH environment has been suggested to induce the DNA fragmentation and apoptosis in tumor cells<sup>11-12</sup>. Duggan et al<sup>13</sup> 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 GSE2144<sup>13</sup> 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 supplemented<sup>14</sup>. Secondly, the complemented data were performed normalization<sup>15</sup>. Next, comparisons were made between control samples and low pH stimulated samples using the LIMMA (Linear Models for Microar-

ray Data) package in R language, to identify the DEGs<sup>16</sup>. Genes only with  $p$ -value  $< 0.05$  and  $\log$  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 clusters<sup>17</sup>. Based on that gene expression is time specific, hierarchical cluster analysis<sup>17</sup> was used to observe the dynamic changes of gene expression at each time point.

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

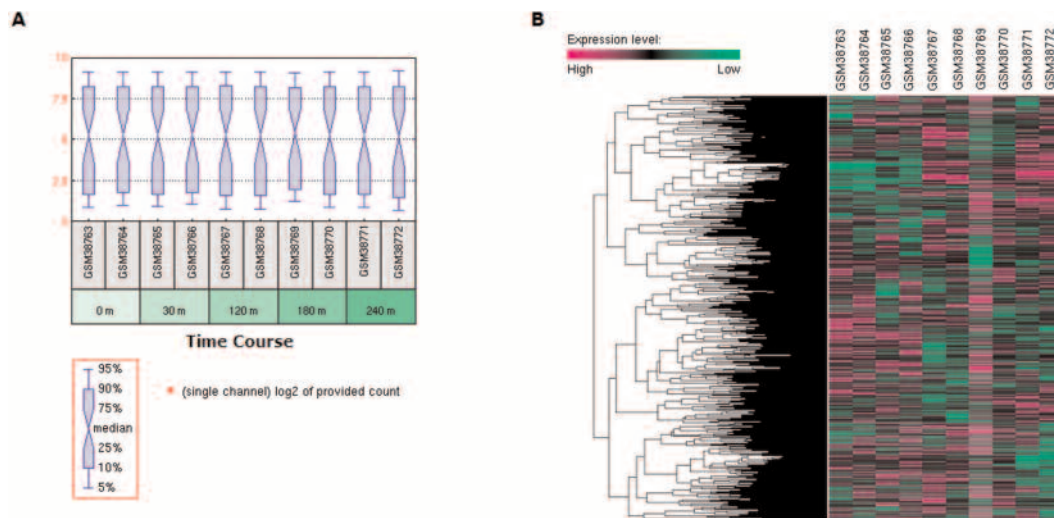
In order to find out the changes of biological functions of SKGT4 cells under the low pH exposure, functional enrichment analysis for the selected DEGs were conducted. DAVID (Database for Annotation, Visualization, and Integrated Discovery) bioinformatics resources consist of an integrated biological knowledge base and analytic tools aimed at systematically extracting biological functions from large gene or protein lists<sup>18</sup>. 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 I.** The results of functional enrichment of DEGs in PPI network.

GO-ID	FDR	Description	Genes	Count
48522	5.1945E-08	Positive regulation of cellular process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	16
48518	1.1575E-07	Positive regulation of biological process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	16
42221	0.000016473	Response to chemical stimulus	EGR1 PIM1 NR4A2 LIF FOS JUN ...	12
32501	0.010568	Multicellular organismal process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	13
50896	0.00014875	Response to stimulus	EGR1 PIM1 NR4A2 LIF FOS JUN ...	15
10556	0.0001055	Regulation of macromolecule biosynthetic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	14
9889	0.00013928	Regulation of biosynthetic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	14
60255	0.00031548	Regulation of macromolecule metabolic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	14
80090	0.00013819	Regulation of primary metabolic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	15
31323	0.00018956	Regulation of cellular metabolic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	15
19222	0.00029354	Regulation of metabolic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	15
19219	0.00047586	Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	13
31326	0.00047915	Regulation of cellular biosynthetic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	13
51171	0.00047985	Regulation of nitrogen compound metabolic process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	13
50794	0.00038703	Regulation of cellular process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	18
50789	0.00070833	Regulation of biological process	EGR1 PIM1 NR4A2 LIF FOS JUN ...	18
65007	0.0014324	Biological regulation	EGR1 PIM1 NR4A2 LIF FOS JUN ...	18
10468	0.00037962	Regulation of gene expression	EGR1 PIM1 NR4A2 LIF FOS JUN ...	13



**Figure 1.** **A**, Expression value of the normalized gene expression; **B**, The hierarchical clustering diagram of gene expression. Red indicated high expression level and green indicated low expression level. The expression value from low to high is showed by gradient from green to red.

imental and predicted interaction information of proteins<sup>19</sup>. In this study, STRING was used to explore the interactions of the common DEGs of the four time points by probabilistic confidence score, which represents a rough estimate of how likely a given association describes a linkage between two proteins<sup>20</sup>. And, then, the significant interactions with confidence score larger than 0.4 were selected to construct PPI network using Cytoscape software<sup>21</sup>.

### Functional Enrichment Analysis of PPI Network

FuncAssociate is a web application that discovers properties enriched in lists of genes or proteins that emerge from large-scale experimentation<sup>22</sup>. In this study, FuncAssociate was used to perform functional annotation of DEGs in the PPI network, based on the hypergeometric distribution algorithm. FDR < 0.05 was chose as cut-off criteria.

## Results

### DEGs Selection

The expression data downloaded from GEO database were firstly normalized, as shown in Figure 1A, and then DEGs at each time point were selected compared with that in the control samples (Figure 1A). The numbers of DEGs of low pH stimulation for 30 min, 120 min, 180 min and 240 min were 285, 365, 277 and 421, respectively.

### 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 lager than 0.9 are shown in Table I, such as EGR1, JUN and FOS.

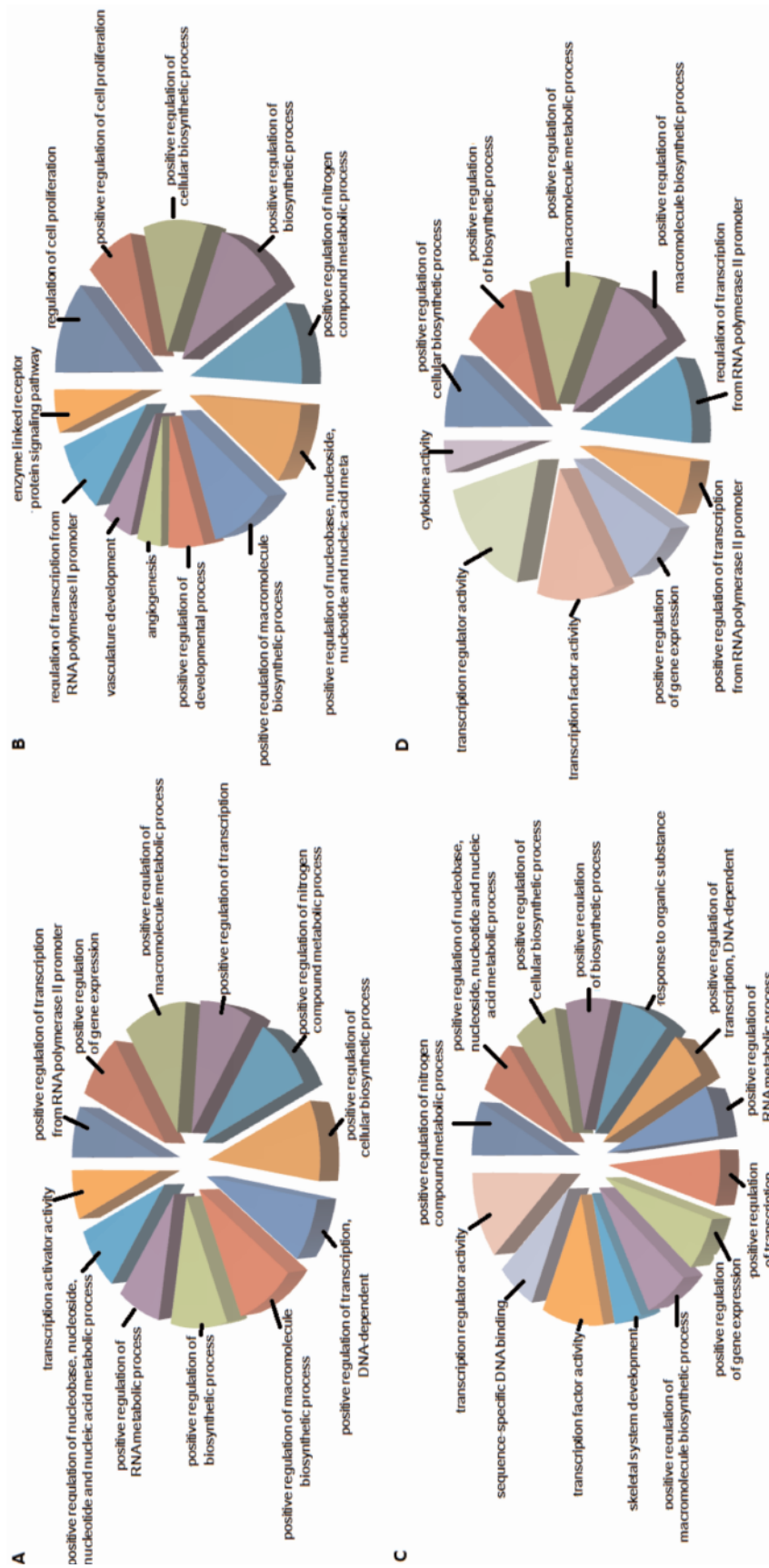


Figure 2. The significant functions of DEGs from 4 time points: A, 30 min; B, 120 min; C, 180 min; and D, 240 min.

**Table II.** Significant interactions of DEGs in the PPI network (confidence score > 0.9).

Node1	Node2	Association score
JUN	ATF3	0.999
JUN	FOS	0.999
LIFR	LIF	0.999
JUN	FOSL1	0.998
FOS	EGR1	0.997
EDN1	FOS	0.995
EDN1	JUN	0.994
JUN	FOSB	0.993
JUN	DUSP1	0.983
FOSB	EGR1	0.976
NR4A1	EGR1	0.976
NR4A2	NR4A1	0.972
FOSB	NR4A1	0.97
FOS	DUSP1	0.969
ATF3	EGR1	0.966
EGR1	DUSP1	0.965
JUN	NR4A1	0.958
JUN	EGR1	0.955
EDN1	CTGF	0.949
NR4A2	FOSB	0.944
FOSB	DUSP1	0.908
MAFF	ATF3	0.9

PPI: protein-protein interaction network.

**Functions Enrichment Analysis of DEGs in PPI Network**

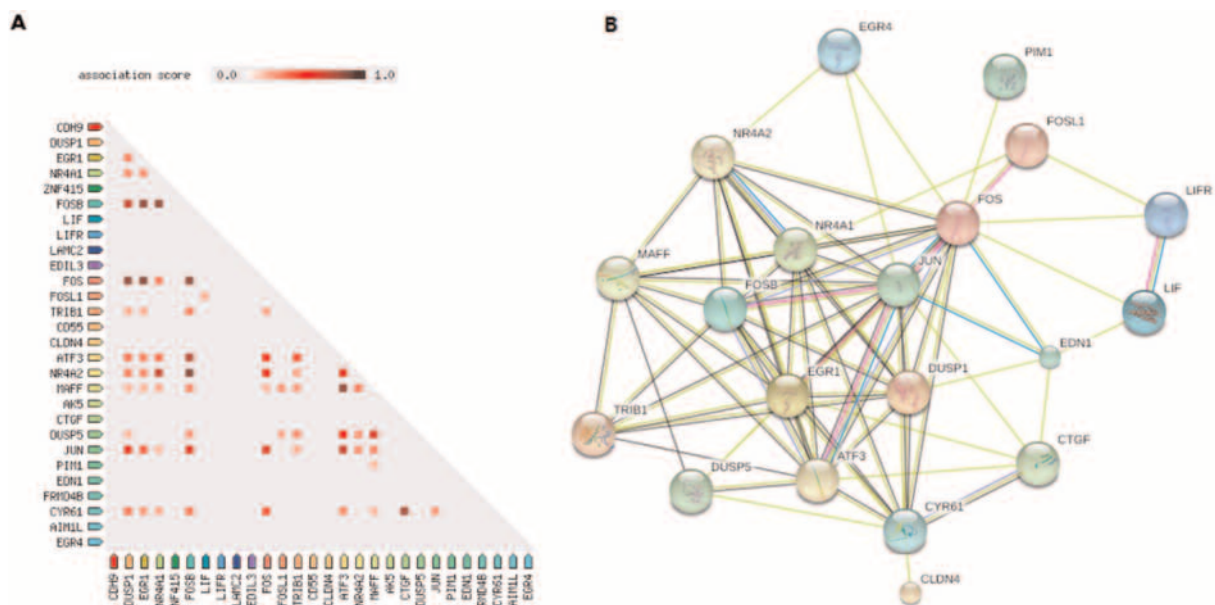
A total of 18 significant GO terms with FDR < 0.05 were enriched (Table II) of the genes in the PPI network. The most significant GO term was

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<sup>23</sup>. 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



**Figure 3.** A, Interactions between common DEGs from 4 time points; B, PPI network of common DEGs.

proteins and as a transcription factor for activates target genes related cell division and mitosis<sup>24</sup>. 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<sup>25-27</sup>. In addition, the EGR1 protein plays a pivotal role in the regulation of cell growth, differentiation and apoptosis. Levkovitz et al<sup>28</sup> have reported that the EGR1 suppresses cerebellar granule cell apoptosis by blocking c-Jun activation. Meanwhile, Chen et al<sup>29</sup> identify EGR1 as a novel target for JUN-induced apoptosis in multiple myeloma. However, Hoffmann et al<sup>30</sup> 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<sup>31-32</sup>, FOS over-expression enhances the motility and invasion of breast and colorectal cancer cells, but inhibits the tumorigenicity of cervical carcinoma cell lines<sup>33-34</sup>. 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<sup>35</sup>. 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 consisted 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<sup>36</sup> 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

- 1) VAKIL N, VAN ZANTEN SV, KAHRILAS P, DENT J, JONES R. The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. *Am J Gastroenterol* 2006; 101: 1900-1920.
- 2) KAHRILAS PJ. Gastroesophageal reflux disease. *N Engl J Med* 2008; 359: 1700-1707.
- 3) KANDULSKI A, MALFERTHEINER P. Gastroesophageal reflux disease-from reflux episodes to mucosal inflammation. *Nat Rev Gastroenterol Hepatol* 2011; 9: 15-22.
- 4) FASS R, OFMAN JJ. Gastroesophageal reflux disease-should we adopt a new conceptual framework& quest. *Am J Gastroenterol* 2002; 97: 1901-1909.
- 5) VAEZI MF, SINGH S, RICHTER JE. Role of acid and duodenogastric reflux in esophageal mucosal injury: a review of animal and human studies. *Gastroenterology* 1995; 108: 1897-1907.
- 6) TURK H, HAUSER B, BRECELJ J, VANDENPLAS Y. Effect of proton pump inhibition on acid, weakly acid and weakly alkaline gastro-esophageal reflux in children. *World J Pediatr* 2013; 9: 36-41.
- 7) IWAKIRI K, KAWAMI N, SANO H, TANAKA Y, UMEZAWA M, KOTOYORI M, HOSHIHARA Y, SAKAMOTO C. Acid and non-acid reflux in Japanese patients with non-erosive reflux disease with persistent reflux symptoms, despite taking a double-dose of proton pump inhibitor: a study using combined pH-impedance monitoring. *J Gastroenterol* 2009; 44: 708-712.
- 8) YOSHIDA N, UCHIYAMA K, KURODA M, SAKUMA K, KOKURA S, ICHIKAWA H, NAITO Y, TAKEMURA T, YOSHIKAWA T, OKANOUE T. Interleukin-8 expression in the esophageal mucosa of patients with gastroesophageal reflux disease. *Scand J Gastroenterol* 2004; 39: 816-822.
- 9) RIEDER F, BIANCANI P, HARNETT K, YERIAN L, FALK GW. Inflammatory mediators in gastroesophageal reflux disease: impact on esophageal motility, fibrosis, and carcinogenesis. *Am J Physiol Gastrointest Liver Physiol* 2010; 298: G571-G581.

- 10) YOSHIDA N. Inflammation and oxidative stress in gastroesophageal reflux disease. *J Clin Biochem Nutr* 2007; 40: 13.
- 11) MIN KH, KIM J-H, BAE SM, SHIN H, KIM MS, PARK S, LEE H, PARK RW, KIM IS, KIM K, KWON IC, JEONG SY, LEE DS. Tumoral acidic pH-responsive MPEG-poly ( $\beta$ -amino ester) polymeric micelles for cancer targeting therapy. *J Control Release* 2010; 144: 259-266.
- 12) ZHANG X, LIN Y, GILLIES RJ. Tumor pH and its measurement. *J Nucl Med* 2010; 51: 1167-1170.
- 13) DUGGAN SP, GALLAGHER WM, FOX EJ, ABDEL-LATIF MM, REYNOLDS JV, KELLEHER D. Low pH induces co-ordinate regulation of gene expression in oesophageal cells. *Carcinogenesis* 2006; 27: 319-327.
- 14) TROYANSKAYA O, CANTOR M, SHERLOCK G, BROWN P, HASTIE T, TIBSHIRANI R, BOTSTEIN D, ALTMAN RB. Missing value estimation methods for DNA microarrays. *Bioinformatics* 2001; 17: 520-525.
- 15) FUJITA A, SATO J, RODRIGUES L, FERREIRA C, SOGAYAR M. Evaluating different methods of microarray data normalization. *BMC Bioinformatics* 2006; 7: 469.
- 16) SMYTH GK. Limma: linear models for microarray data. *Bioinformatics and computational biology solutions using R and Bioconductor*. Springer 2005; 397-420.
- 17) EISEN MB, SPELLMAN PT, BROWN PO, BOTSTEIN D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998; 95: 14863-14868.
- 18) DA WEI HUANG BTS, LEMPICKI RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2008; 4: 44-57.
- 19) SZKLARCZYK D, FRANCESCHINI A, KUHN M, SIMONOVIC M, ROTH A, MINGUEZ P, DOERKS T, STARK M, MULLER J, BORK P, JENSEN LJ, VON MERING C. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* 2011; 39: D561-D568.
- 20) VON MERING C, JENSEN LJ, SNEL B, HOOPER SD, KRUPP M, FOGLEIRINI M, JOUFFRE N, HUYNEN MA, BORK P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res* 2005; 33: D433-D437.
- 21) VLASBLOM J, WU S, PU S, SUPERINA M, LIU G, ORSI C, WODAK SJ. GenePro: a Cytoscape plug-in for advanced visualization and analysis of interaction networks. *Bioinformatics* 2006; 22: 2178-2179.
- 22) BERRIZ GF, BEAVER JE, CENIK C, TASAN M, ROTH FP. Next generation software for functional trend analysis. *Bioinformatics* 2009; 25: 3043-3044.
- 23) DANJO A, SAKATA H. Dysphagia associated with gastroesophageal reflux disease is improved by proton pump inhibitor. *Dig Dis Sci* 2005; 50: 1921-1926.
- 24) PATTON WF. Fluorescence detection of glycoproteins in gels and on electroblots. *Curr Protoc Cell Biol* 2002; 6.8. 1-6.8. 15.
- 25) WU MY, ZHUANG CX, YANG HX, LIANG Y-R. Expression of Egr-1, c-fos and cyclin D1 in esophageal cancer and its precursors: an immunohistochemical and in situ hybridization study. *World J Gastroenterol* 2004; 10: 476-480.
- 26) LEE S-H, BAHN JH, CHOI CK, WHITLOCK NC, ENGLISH AE, SAFE S, BAEK SJ. ESE-1/EGR-1 pathway plays a role in tolfenamic acid-induced apoptosis in colorectal cancer cells. *Mol Cancer Ther* 2008; 7: 3739-3750.
- 27) LIU J, LIU Y-G, HUANG R, YAO C, LI S, YANG W, YANG D, HUANG RP. Concurrent down-regulation of Egr-1 and gelsolin in the majority of human breast cancer cells. *Cancer Genomics Proteomics* 2007; 4: 377-385.
- 28) LEVKOVITZ Y, BARABAN JM. A dominant negative inhibitor of the Egr family of transcription regulatory factors suppresses cerebellar granule cell apoptosis by blocking c-Jun activation. *J Neurosci* 2001; 21: 5893-5901.
- 29) CHEN L, WANG S, ZHOU Y, WU X, ENTIN I, EPSTEIN J, YACCOBY S, XIONG W, BARLOGIE B, SHAUGHNESSY JD JR, ZHAN F. Identification of early growth response protein 1 (EGR-1) as a novel target for JUN-induced apoptosis in multiple myeloma. *Blood* 2010; 115: 61-70.
- 30) HOFFMANN E, ASHOURI J, WOLTER S, DOERRIE A, DITTRICH-BREIHZOL O, SCHNEIDER H, WAGNER EF, TROPFMAIR J, MACKMAN N, KRACHT M. Transcriptional regulation of EGR-1 by the interleukin-1-JNK-MKK7-c-Jun pathway. *J Biol Chem* 2008; 283: 12120-12128.
- 31) DURCHDEWALD M, ANGEL P, HESS J. The transcription factor Fos: a Janus-type regulator in health and disease. *Histol Histopathol* 2009; 24: 1451.
- 32) EFERL R, WAGNER EF. AP-1: a double-edged sword in tumorigenesis. *Nat Rev Cancer* 2003; 3: 859-868.
- 33) KUSTIKOVA O, KRAMEROV D, GRIGORIAN M, BEREZIN V, BOCK E, LUKANIDIN E, TULCHINSKY E. Fra-1 induces morphological transformation and increases in vitro invasiveness and motility of epithelioid adenocarcinoma cells. *Mol Cell Biol* 1998; 18: 7095-7105.
- 34) TKACH V, TULCHINSKY E, LUKANIDIN E, VINSON C, BOCK E, BEREZIN V. Role of the Fos family members, c-Fos, Fra-1 and Fra-2, in the regulation of cell motility. *Oncogene* 2003; 22: 5045-5054.
- 35) LAMPH WW, WAMSLEY P, SASSONE-CORSI P, VERMA IM. Induction of proto-oncogene JUN/AP-1 by serum and TPA. *Nature* 1988; 334: 629-631.
- 36) McMAHON AP, CHAMPION J, McMAHON J, SUKHATME V. Developmental expression of the putative transcription factor Egr-1 suggests that Egr-1 and c-fos are coregulated in some tissues. *Development* 1990; 108: 281-287.