RESEARCH COMMUNICATION

Transcriptome Network Analysis Reveals Potential Candidate Genes for Esophageal Squamous Cell Carcinoma

Zheng Ma[1], Wei Guo[1], Hui-Jun Niu[1], Fan Yang[1], Ru-Wen Wang[1], Yao-Guang Jiang[1], Yun-Ping Zhao[1]*

Abstract

   The esophageal squamous cell carcinoma (ESCC) is an aggressive tumor with a poor prognosis. Understanding molecular changes in ESCC should improve identification of risk factors with different molecular subtypes and provide potential targets for early detection and therapy. Our study aimed to obtain a molecular signature of ESCC through the regulation network based on differentially expressed genes (DEGs). We used the GSE23400 series to identify potential genes related to ESCC. Based on bioinformatics we constructed a regulation network. From the results, we could establish that many transcription factors and pathways closely related with ESCC were linked by our method. STAT1 also arose as a hub node in our transcriptome network, along with some transcription factors like CCNB1, TAP1, RARG and IFITM1 proven to be related with ESCC by previous studies. In conclusion, our regulation network provided information on important genes which might be useful in investigating the complex interacting mechanisms underlying the disease.

Keywords: Esophageal squamous cell carcinoma - network - differentially expressed genes

Introduction

   Esophageal cancers (or oesophageal cancer) which arises from the epithelium, or surface lining of the esophagus are typically carcinomas. ESCC fall into two classes. One subtype is primarily squamous cell cancer, arising from the cells that line the upper part of the esophagus and adenocarcinoma, and this cancer is similar to head and neck cancer in their appearance and associates with tobacco and alcohol consumption. The other is adenocarcinomas, which are often associated with a history of gastroesophageal reflux disease and Barrett's esophagus, arising from glandular cells that are present at the junction of the esophagus and stomach (Holmes and Vaughan, 2007).

   Recently, many researches are carried out to understand the precise molecular mechanisms of esophageal squamous cell carcinoma (ESCC), which could be summarized as three aspects: tumour suppressor genes, oncogenes and apoptotic genes. P53 is a tumour suppressor that halts progression in both the G 1 and G 2 phase of the cell cycle to assess DNA damage. Accumulation of p53 in the normal oesophagus, suggested that the loss of suppressor function p53 might be an early event in carcinogenesis of the oesophagus. Mutations in codons 175, 248 and 273 of p53 are considered to have growth advantages to progress to invasive squamous cell carcinoma and

occured most frequently. LOH of the Rb gene was found correlated with the loss of pRb protein expression and associated with p53 alterations in human oesophageal cancer. It is suggested that associated Rb and p53 inactivation may be the major event in the development and progression of oesophageal cancer, due to the greater selective advantage of the affected cells. The oncogenes activated most frequent in oesophageal cancer are cyclin D1, c-erbB1 and 2, c-myc, c-ras, Int-2/hst-1, and EGFR. Cyclin D1 gene amplification was found in about 32% of oesophageal tumours, and 92% of these tumours showed cyclin D1 overexpression. Frat1 expression was also found to be relatively high in human oesophageal cancer cell lines through the activation of the WNT-h-catenin-TCF signaling pathway. Proteins that have been found to be aberrantly expressed within oesophageal cancer including the anti-apoptotic proteins Bcl-2 (up-regulated) and pro-apoptotic protein Bax (down-regulated) (Kuwano et al., 2005; McCabe and Dlamini, 2005).

   DNA microarray analysis as a global approach is applied to investigate physiological mechanisms in health and disease (Spies et al., 2002). The microarray experiment significant reduction of MDR1 expression was observed in patients suggesting that MDR1 might affect the sensitivity of 5FU and/or CDDP in the adjuvant chemotherapy for esophageal cancers (Kihara et al., 2001). Genomic expression profiling evolves as a useful tool to

Department of General Thoracic Surgery, Daping Hospital and Institute of Surgery Research, The Third Millitary Medical University, Chongqing, China *For correspondence: zhaoyunpingzh@gmail.com

Table 1. Regulation Data from TRANSFAC and TRED

Source	Regulation	TFs	Targets	Link
TRANSFAC	774	219	265	http://www.gene-regulation.com/pub/databases.html
TRED	5722	102	2920	http://rulai.cshl.edu/TRED/
total	6328	276	3002	

identify novel pathomechanisms in human cancer (Guo, 2003).

The purpose of this paper is to propose a hypothesis that a transcriptome network can be developed such that a set of transcription factors, regulated the differently expression genes are induced by ESCC and can be identified and modulated. Further analysis of the genes and pathways in the network were taken to identify potential mechanisms responded to the ESCC. The study does not address regulation network but searches for the significance pathways related to ESCC.

Materials and Methods

Data Source

Affymetrix microarray data: One transcription profiles of Ischemic cardiomyopathy GSE23400 (Su et al., 2011) were obtained from a public functional genomics data repository GEO (http://www.ncbi.nlm.nih.gov/geo/). Only 106 chips are usable. Samples derived from 53 DCMi and 53 healthy control patients were hybridised onto Affymetrix Human Genome U133A Array (Wachi et al., 2005).

Pathway data: KEGG (Kyoto Encyclopedia of Genes and Genomes) is a collection of online database dealing with genomes, enzymatic pathways and biological chemicals (Kanehisa, 2002). The PATHWAY database records networks of molecular interactions in the cells, and variants of them specific to particular organisms (http://www.genome.jp/kegg/). Total 130 pathways, involving 2,287 genes, were collected from KEGG.

Regulation data: There are approximately 2600 proteins in the human genome that contain DNA-binding domains, and most of which are presumed to function as transcription factors (Wachi et al., 2005). The combinatorial use of a subset of the approximately 2000 human transcription factors easily accounts for the unique regulation of each gene in the human genome during development (Brivanlou and Darnell, 2002).

These transcription factors are grouped into 5 super class families, based on the presence of conserved DNA-binding domains. TRANSFAC database contains data on transcription factors, and their experimentally-proven binding sites and regulated genes (Wingender, 2008).

Transcriptional Regulatory Element Database (TRED) has been built to increase the needs of an integrated repository for both cis- and trans- regulatory elements in mammals (Jiang et al., 2007). TRED has the curation for transcriptional regulation information, including transcription factor binding motifs and experimental evidence. The curation is currently focusing on target genes of 36 cancer-related TF families.

774 pairs of regulatory relationship between 219 transcription factors (TFs) and 265 target genes were collected from TRANSFAC (http://www.gene-regulation.com/pub/databases.html).

5,722 pairs of regulatory relationship between 102 transcription factors (TFs) and 2,920 target genes were collected from TRED (http://rulai.cshl.edu/TRED/).

Combined the two regulation datasets, total 6328 regulatory relationships between 276 TFs and 3,002 target genes were collected (Table 1).

miRNA datasets: Integrate miRNA to disease datasets and miRNA to target genes datasets to find the relation between disease and genes. A human miRNA disease database (HMDD) (Lu et al., 2008) which manually retrieves the associations of miRNA and disease from literatures contains 444 miRNA genes, 259 diseases, 1,149 publications, and 2,886 miRNA-disease associations. miR2Disease (Jiang et al., 2009) provides a comprehensive resource of miRNA deregulation in various human diseases. Total 349 miRNA and 163 diseases were collected. Merge the above two database, 5,036 relationships were collected. We also integrate starbase (Yang et al., 2011), miR2Disease, miRecords (Xiao et al., 2009) and TarBase (Papadopoulos et al., 2009) database of miRNA-target genes, whether is proved by experiment. At last, 211,464 relations between miRNA and genes were selected.

Methods

Differentially expressed genes (DEGs) analysis: For the GSE23400 dataset, the limma method (Smyth, 2004) was used to identify DEGs. The original expression datasets from all conditions were processed into expression estimates using the RMA method with the default settings implemented in Bioconductor, and then constructed the linear model. The DEGs only with the fold change value larger than 1.5 and p-value less than 0.05 were selected.

Co-expression analysis: To demonstrate the potential regulatory relationship, the Pearson Correlation Coefficient (PCC) was calculated for all pair-wise comparisons of gene-expression values between TFs and the DEGs. The regulatory relationships whose absolute PCC are larger than 0.6 were considered as significant.

Gene Ontology analysis: The Biological Networks Gene Ontology tool (BiNGO) is an open-source Java tool to determine which Gene Ontology (GO) terms are significantly overrepresented in a set of genes. The BiNGO analysis (Maere et al., 2005) was used to identify over-represented GO categories in biological process.

Regulation network construction: Using the regulation data that have been collected from TRANSFAC database and TRED database, we matched the relationships between differentially expressed TFs and its differentially expressed target genes.

Base on the above two regulation datasets and the pathway relationships of the target genes, we constructed the regulation networks by Cytoscape (Shannon et al.,

2003). Base on the significant relationships (PCC > 0.6 or PCC < -0.6) between TFs and its target genes, 33 putative regulatory relationships were predicted between 7 TFs and 22 target genes.

Significance analysis of pathway: We adopted an impact analysis that includes the statistical significance of the set of pathway genes, and also considered other crucial factors such as the magnitude of each gene's expression change, the topology of the signaling pathway, their interactions, etc (Draghici et al., 2007). In this model, the Impact Factor (IF) of a pathway P_i is calculated as the sum of two terms:

$$IF(Pi) = \log(\frac{1}{pi}) + \frac{\sum_{g \square Pi} |PF(g)|}{|\triangle E| \times N_{de}(Pi)}$$

The first term is a probabilistic term that captures the significance of the given pathway P_i from the perspective of the set of genes contained in it.

It is obtained by using the hyper geometric model in which pi is the probability of obtaining at least the observed number of differentially expressed gene, Nde, just by chance (Tavazoie et al., 1999; Draghici et al., 2003).

The second term is a functional term that depends on the identity of the specific genes that are differentially expressed as well as on the interactions described by the pathway (i.e., its topology).

The second term sums up the absolute values of the perturbation factors (PFs) for all genes g on the given pathway P_i.

The PF of a gene g is calculated as follows:

$$PF(g) = \triangle E(g) + \sum_{u \square USg} b_{ug} \times \frac{PF(u)}{N_{ds}(u)}$$

In this equation, the first term ΔE (g) captures the quantitative information measured in the gene expression experiment. The factor ΔE (g) represents the normalized measured expression change of the gene g. The first term ΔE (g) in the above equation is a sum of all PFs of the genes u directly upstream of the target gene g, normalized by the number of downstream genes of each such gene N_{ds} (u), and weighted by a factor βug, which reflects the type of interaction: $β_{ug} = 1$ for induction, $β_{ug} = -1$ for repression (KEGG supply this information about the type of interaction of two genes in the description of the pathway topology). US_g is the set of all such genes upstream of g. We need to normalize with respect to the size of the pathway by dividing the total perturbation by the number of differentially expressed genes on the given pathway, N_{de} (P_i). In order to make the IFs as independent as possible from the technology, and also comparable between problems, we also divide the second term in equation 1 by the mean absolute fold change ΔE, and calculated across all differentially expressed genes. The results of the significance analysis of pathway were shown in Table 3.

Regulation network between TFs and pathways: To further investigate the regulatory relationships between TFs and pathways, we mapped DEGs to pathways and got a regulation network between TFs and pathways (Figure 2).

miRNA network construction: Using the miRNA-disease data that have been collected from HMDD and miR2Disease database, we matched the disease only to the esophageal. The esophageal related miRNA selected and then mapped the miRNA to the target genes. At last, 14 miRNA and 431 DEGs were used to construct the network.

Results

Regulation network construction in ESCC

To get pathway-related DEGs of ESCC, we obtained publicly available microarray data sets GSE23400 from GEO. After microarray analysis, the differentially expressed genes with the fold change value larger than 1.5 of GSE23400 and p-value less than 0.05 were selected. 2008 genes were selected as DEGs from GSE23400. To get the regulatory relationships, the co-expressed value (PCC ≥0.6) was chosen as the threshold. Finally, we got only 17 regulatory relationships between 8 TFs and their 16 differently expressed target genes. By integrating the regulatory relationships above, a regulation network of ESCC was built between TFs and its target genes (Figure 1). In this network, STAT1 with the higher degree forms a local network which suggesting it may plays an important role in ESCC. Both STAT1 and IRF9 regulate the ISG15 target genes directly. Besides, SP1 regulating 3 DEGs and TFAP2A regulating the CCNB1 were also observed in our network.

GO analysis of the regulation network in ESCC

Several Gene Ontology (GO) categories were enriched among these genes in the regulatory network, including negative regulation of biological process, tissue development, embryonic eye morphogenesis and so on (Table 2).

Significant pathway in ESCC

To identify the relevant pathways changed in lung adenocarcinoma, we used a statistical approach on pathway level. Significance analysis at single gene level may suffer from the limited number of samples and experimental noise that can severely limit the power of the
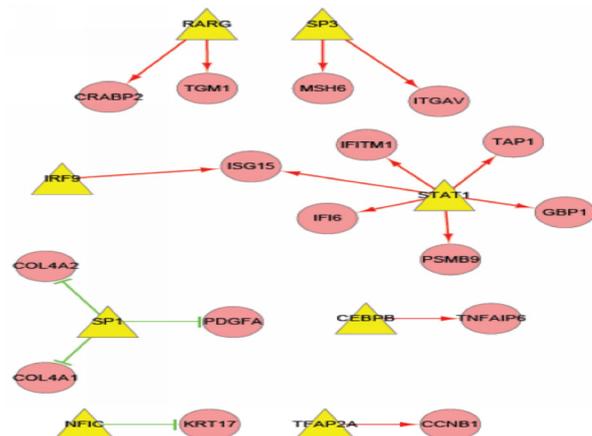

Figure 1. Regulation Network Construction in ESCC

Table 2. GO Biological Process Analysis

GO-ID	Description	count	p-value	corr p-value
48519	negative regulation of biological process	14	2.80E-07	2.32E-04
9888	tissue development	9	1.20E-06	4.97E-04
48048	embryonic eye morphogenesis	3	2.43E-06	6.72E-04
48513	organ development	12	5.14E-06	1.06E-03
48523	negative regulation of cellular process	12	6.95E-06	1.15E-03
44419	interspecies interaction between organisms	6	9.91E-06	1.36E-03
7398	ectoderm development	5	1.46E-05	1.36E-03
30219	megakaryocyte differentiation	2	1.48E-05	1.36E-03
51704	multi-organism process	8	1.85E-05	1.36E-03
31324	negative regulation of cellular metabolic process	8	1.88E-05	1.36E-03

Table 3. Pathway Significant Analysis

Database Name	Pathway Name	Impact Factor	%PathwayGenes in Input	corrected gamma p-value
KEGG	Cardiac muscle contraction	23.687	8.046	1.27E-09
KEGG	Cell cycle	21.205	31.356	1.37E-08
KEGG	ECM-receptor interaction	17.348	34.524	5.36E-07
KEGG	DNA replication	14.992	44.444	4.93E-06
KEGG	Focal adhesion	14.383	21.182	8.72E-06
KEGG	Proteasome	9.932	31.25	5.31E-04
KEGG	p53 signaling pathway	9.397	27.536	8.63E-04
KEGG	Graft-versus-host disease	9.074	2.381	0.001155
KEGG	Gap junction	8.384	9.375	0.002144
KEGG	Biosynthesis of unsaturated fatty acids	8.239	40.909	0.00244

Figure 2. Regulation Network of TF-PATHWAY

Figure 3. miRNA Network. The red points strand for the miRNA, the yellow points strand for the TFs and the pink points strand for target genes
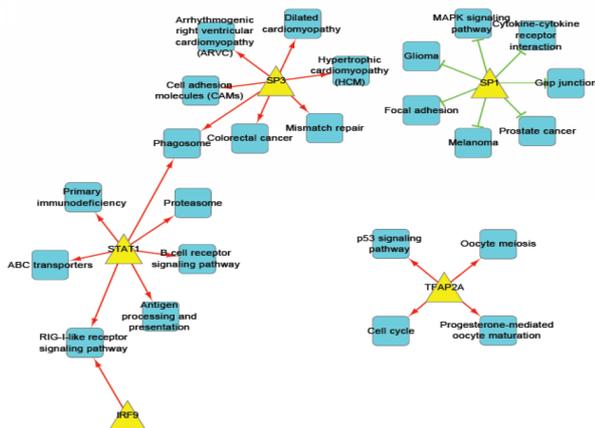
chosen statistical test. Pathway can provide an alternative way to relax the significance threshold applied to single genes and may lead to a better biological interpretation. So, we adopted a pathway based impact analysis method that contained many factors including the statistical significance of the set of differentially expressed genes in the pathway, the magnitude of each gene's expression change, the topology of the signaling pathway, and their interactions and so on. The impact analysis method yields many significant pathways contained Cardiac muscle contraction, Cell cycle, ECM-receptor interaction and so on (Table 3).

Regulation network between TFs and pathways in ESCC

To further investigate the regulatory relationships between TFs and pathways, we mapped DEGs to pathways and got a regulation network between TFs and pathways (Figure 2). In the network, SP3 and ATAT1 were shown as hub nodes linked to lots of esophageal related pathways. Some of TFs Interactive regulated lots of pathways, such as STAT1 and SP3 both regulated the Phagosome
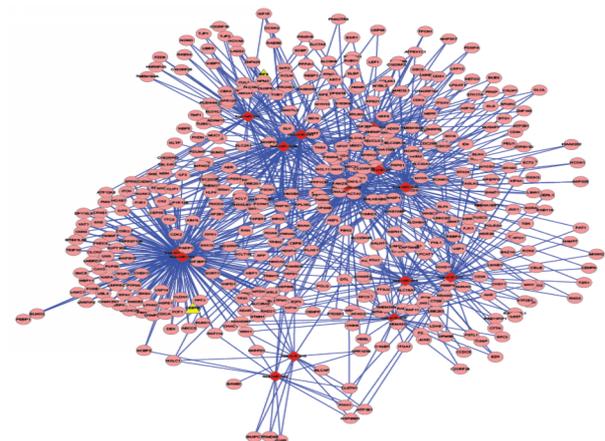
pathway, and STAT1 and IRF9 both regulated the RIG-I-like receptor signaling pathway.

miRNA network analysis

The miRNA network further confirms that the DEGs were highly related to ESCC. Finally, we selected 14 miRNA and 431 DEGs that response to esophageal (Figure 3).

Discussion

From the result of regulation network construction in ESCC, we could find that many TFs and pathways closely related with ESCC have been linked by our method. The gene STAT1 is shown as hub nodes in our transcriptome network. Although the role of STAT1, Sp1 and ISG15 in ESCC has not been investigated to verify, some evidences also suggest that they may play important roles in response to ESCC.

Sp1 transcription factor was over-expressed in ESCC cells and the increased Sp1 staining was observed in esophageal tumors from patients (Papineni et al., 2009). However, Sp1 plays an important role in ESCC through regulating other target genes expression, such as fascin. Fascin expression was enhanced by Sp1 over-expression and blocked by Sp1 RNAi knockdown. Specific inhibition of ERK1/2 decreased phosphorylation levels of Sp1, and thus suppressed the transcription of the FSCN1, resulting in the down-regulation of fascin. Stimulation with EGF could activate the ERK1/2 pathway and increase phosphorylation levels of Sp1 to enhance fascin expression (Lu et al., 2010).

STAT1 can be phosphorylated by the receptor associated kinases, and translocate to the cell nucleus where they act as transcription activators in response to cytokines and growth factors such as EGF. EGF treatment leads to a strong growth arrest in three esophageal squamous cell carcinoma cell lines. STAT1 was also found to be activated by EGF in this three cell lines. Therefore, the EGF-STAT1 pathway may be intrinsic to esophageal epithelial lineage of cells and is lost in a considerable fraction of the carcinomas (Watanabe et al., 2001). And, the introduction of a dominant-negative STAT1 construct into TE8 cells abolished not only ESCC cells growth inhibition but also p21/WAF1 induction by EGF (Ichiba et al., 2002).

RARG (retinoic acid receptor gamma 1), is a nuclear hormone receptor that act as ligand-dependent transcriptional regulators. RARs have numerous target genes, which have retinoic response elements in their promoter regions. RARG were down-regulated in esophageal adenocarcinoma tissues (Hourihan et al., 2003).

Cyclin B1 plays an important role as a mitotic cyclin in the G (2)-M phase transition during the cell cycle. One-, 3-, and 5-year survival rates in esophageal squamous cell carcinoma (ESCC) cells with cyclin B1 expression were significantly lower than those in ESCC cells without cyclin B1 expression (Nozoe et al., 2002). And ESCC cells over-expressing cyclin B1 reveal strong invasive growth and high potential of metastasis to lung in xenograft mice. Suppression of cyclin B1 expression via small interfering RNA approach specifically inhibits their ability to metastasize (Song et al., 2008).

TAP1 protein is a member of the MDR/TAP subfamily of ATP-binding cassette (ABC) transporters. This protein is involved in the pumping of degraded cytosolic peptides across the endoplasmic reticulum into the membrane-bound compartment where ABC1 molecules assemble. TAP1, as an antigen-processing machinery component, was down-regulated in 44% of ESCC lesions. TAP1 loss was significantly associated with tumor grading, lymph node metastasis and depth of invasion (Ayshamgul et al., 2011).

ISG15 is an ubiquitin-like protein that becomes conjugated to many cellular proteins upon activation by interferon-alpha and -beta. 5-Fluorouracil (FU) is a chemotherapeutic agent commonly used against ESCC. Study showed that three IFN-related genes, an IFN receptor gene (IFNAR2) and two IFN-stimulated genes (ISG15K, ISG-54K) were up-regulated after three cell lines (T.T, TE-2, and TE-6), derived from human ESCC, were exposed to 5-FU indicating a combination of 5-FU and an IFN, may be particularly efficacious in ESCC (Matsumura et al., 2005).

Prior laboratory prediction of individual drug response is of key importance in ESCC. Interferon induced transmembrane protein 1 (IFITM1) gene alone, being suggested as a key gene of Wnt pathway, was commonly selected in 5-fluorouracil (5-FU) and cis-platinum (CDDP) screening methods using a new statistical analysis of oligonucleotide microarray expression data, based on a two-dimensional mixed normal model. The results suggested the IFITM1 gene was a novel critical biomarker of CDDP response in ESCC (Fumoto et al., 2008). IFITM1 was up-regulated in esophageal tumor tissue from patients (Chattopadhyay et al., 2009).

Using a case-control design, the authors measured the odds of being diagnosed with colorectal adenocarcinoma some time in life among persons diagnosed with adenocarcinoma of the esophagus relative to persons diagnosed with squamous cell carcinomas of the esophagus. The results concluded that men with esophageal adenocarcinoma may be more likely to be diagnosed with colorectal cancer in their lifetime than expected, but the opposite association may exist for women providing additional evidence that some colorectal and esophageal adenocarcinoma share a common etiology (Vaughan et al., 1995). Some studies also reported the risk of colorectal cancer in Barrett's esophagus patients was more than six times greater than in those patients participating in colonic screening programs (Morgan, 1996).

Primary malignant melanoma of the esophagus is a rare but aggressive tumor that accounts for <1% of all ESCC (Lin et al., 2006). To date, dozens of melanoma-associated antigens (MAGEs) have been identified in relation with tumor genesis which can be recognized by T lymphocytes to induce immune reaction. Therefore, MAGE genes take part in the immune process by targeting some early tumor cells for immune destruction. Over-expression of MAGE-A1, 2b, 3, 4, 6, 9, 10, and 12 was found in esophageal adenocarcinoma relative to Barrett's metaplasia which may provide potential targets for immunotherapy in patients with esophageal adenocarcinoma (Lin et al., 2004).

One case of ESCC revealed serum prostate-specific antigen (PSA) and γ-seminoprotein levels elevated in blood biochemistry detection. Furthermore, PSA-positive mRNA was demonstrated in the tissue of the esophageal tumor by RT-PCR. Previous reports also suggested esophagus cancer might result from prostate cancer metastasis (Nakamura et al., 1997). Annexin I is a pleotrophic, calcium-dependent phospholipid binding protein. Study indicated loss of inhibition of annexin I appeared associated with a lack of cellular differentiation, and annexin I protein expression was decreased in human esophageal and prostate cancer, which indirectly suggested that some correlation consisted in esophageal and prostate carcinoma onset (Paweletz et al., 2000).

A deeper understanding of transcription factors and their regulated genes remain an area of intense research

activity in futures. Our regulation network is useful in investigating the complex interacting mechanisms of transcription factors and their regulated genes in disease. However, further experiments are still needed to confirm the conclusion. Our network could explain the mechanism of esophageal not only from the regulation network but the miRNA network. The 14 miRNAs were related to esophageal collected from HMDD and miR2Disease database.

Acknowledgements

References

Ayshamgul H, Ma H, Ilyar S, et al (2011). Association of defective HLA-I expression with antigen processing machinery and their association with clinicopathological characteristics in Kazak patients with esophageal cancer. *Chin Med J (Engl)*, **124**, 341-6.

Brivanlou AH, Darnell JE, Jr (2002). Signal transduction and the control of gene expression. *Science*, **295**, 813-8.

Chattopadhyay I, Phukan R, Singh A, et al (2009). Molecular profiling to identify molecular mechanism in esophageal cancer with familial clustering. *Oncol Rep*, **21**, 1135-46.

Draghici S, Khatri P, Martins RP, et al (2003). Global functional profiling of gene expression. *Genomics*, **81**, 98-104.

Draghici S, Khatri P, Tarca AL, et al (2007). A systems biology approach for pathway level analysis. *Genome Res*, **17**, 1537-45.

Fumoto S, Shimokuni T, Tanimoto K, et al (2008). Selection of a novel drug-response predictor in esophageal cancer: a novel screening method using microarray and identification of IFITM1 as a potent marker gene of CDDP response. *Int J Oncol*, **32**, 413-23.

Guo QM (2003). DNA microarray and cancer. *Curr Opin Oncol*, **15**, 36-43.

Holmes RS, Vaughan TL (2007). Epidemiology and pathogenesis of esophageal cancer. *Semin Radiat Oncol*, **17**, 2-9.

Hourihan RN, O'Sullivan GC, Morgan JG (2003). Transcriptional gene expression profiles of oesophageal adenocarcinoma and normal oesophageal tissues. *Anticancer Res*, **23**, 161-5.

Ichiba M, Miyazaki Y, Kitamura S, et al (2002). Epidermal growth factor inhibits the growth of TE8 esophageal cancer cells through the activation of STAT1. *J Gastroenterol*, **37**, 497-503.

Jiang C, Xuan Z, Zhao F, Zhang MQ (2007). TRED: a transcriptional regulatory element database, new entries and other development. *Nucleic Acids Res*, **35**, D137-40.

Jiang Q, Wang Y, Hao Y, et al (2009). miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Re*, **37**, D98-104.

Kanehisa M (2002). The KEGG database. *Novartis Found Symp*, **247**, 91-101.

Kihara C, Tsunoda T, Tanaka T, et al (2001). Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles. *Cancer Res*, **61**, 6474-9.

Kuwano H, Kato H, Miyazaki T, et al (2005). Genetic alterations in esophageal cancer. *Surgery today*, **35**, 7-18.

Lin J, Lin L, Thomas DG, et al (2005). Melanoma-associated antigens in esophageal adenocarcinoma: identification of novel MAGE-A10 splice variants. *Clin Cancer Res*, **10**,

5708-16.

Lin PW, Lee RC, Chern MS, et al (2006). Primary malignant melanoma of the esophagus. *J Chinese Med Association*, **69**, 334-7.

Lu M, Zhang Q, Deng M, et al (2008). An analysis of human microRNA and disease associations. *PLoS One*, **3**, e3420.

Lu XF, Li EM, Du ZP, et al (2010). Specificity protein 1 regulates fascin expression in esophageal squamous cell carcinoma as the result of the epidermal growth factor/extracellular signal-regulated kinase signaling pathway activation. *Cell Mol Life Sci*, **67**, 3313-29.

Maere S, Heymans K, Kuiper M (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **21**, 3448-9.

Matsumura Y, Yashiro M, Ohira M, et al (2005). 5-Fluorouracil up-regulates interferon pathway gene expression in esophageal cancer cells. *Anticancer Res*, **25**, 3271-8.

McCabe M, Dlamini Z (2005). The molecular mechanisms of oesophageal cancer. *Int Immunopharmacol*, **5**, 1113-30.

Morgan G (1996). Relationship between colorectal and esophageal cancer. *Dis Colon Rectum*, **39**, 237.

Nakamura T, Mohri H, Shimazaki M, et al (1997). Esophageal metastasis from prostate cancer: diagnostic use of reverse transcriptase-polymerase chain reaction for prostate-specific antigen. *J Gastroenterol*, **32**, 236-40.

Nozoe T, Korenaga D, Kabashima A, et al (2002). Significance of cyclin B1 expression as an independent prognostic indicator of patients with squamous cell carcinoma of the esophagus. *Clin Cancer Res*, **8**, 817-22.

Papadopoulos GL, Reczko M, Simossis VA, et al (2009). The database of experimentally supported targets: a functional update of TarBase. *Nucleic Acids Res*, **37**, D155-8.

Papineni S, Chintharlapalli S, Abdelrahim M, et al (2009). Tolfenamic acid inhibits esophageal cancer through repression of specificity proteins and c-Met. *Carcinogenesis*, **30**, 1193-201.

Paweletz CP, Ornstein DK, Roth MJ, et al (2000). Loss of annexin 1 correlates with early onset of tumorigenesis in esophageal and prostate carcinoma. *Cancer Res*, **60**, 6293.

Shannon P, Markiel A, Ozier O, et al (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, **13**, 2498-504.

Smyth GK (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*, **3**, Article 3.

Song Y, Zhao C, Dong L, et al (2008). Overexpression of cyclin B1 in human esophageal squamous cell carcinoma cells induces tumor cell invasive growth and metastasis. *Carcinogenesis*, **29**, 307-15.

Spies M, Dasu MR, Svrakic N, et al (2002). Gene expression analysis in burn wounds of rats. *Am J Physiol Regul Integr Comp Physiol*, **283**, R918-30.

Su H, Hu N, Yang HH, et al (2011). Global gene expression profiling and validation in esophageal squamous cell carcinoma and its association with clinical phenotypes. *Clin Cancer Res*, **17**, 2955-66.

Tavazoie S, Hughes JD, Campbell MJ, et al (1999). Systematic determination of genetic network architecture. *Nat Genet*, **22**, 281-5.

Vaughan TL, Kiemeney L, McKnight B (1995). Colorectal cancer in patients with esophageal adenocarcinoma. *Cancer Epidemiol Biomarkers Prev*, **4**, 93.

Wachi S, Yoneda K, Wu R (2005). Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics*, **21**, 4205-8.

Watanabe G, Kaganoi J, Imamura M, et al (2001). Progression

of esophageal carcinoma by loss of EGF-STAT1 pathway. *Cancer J*, **7**, 132-9.

Wingender E (2008). The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Brief Bioinform*, **9**, 326-32.

Xiao F, Zuo Z, Cai G, et al (2009). miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res*, **37**, D105-10.

Yang JH, Li JH, Shao P, et al (2011). starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. *Nucleic Acids Res*, **39**, D202-9.